Office de la Propriété Intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An agency of Industry Canada

CA 2355255 A1 2000/06/22

(21) 2 355 255

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 1999/12/17
(87) Date publication PCT/PCT Publication Date: 2000/06/22
(85) Entrée phase nationale/National Entry: 2001/06/13
(86) N° demande PCT/PCT Application No.: FR 99/03179
(87) N° publication PCT/PCT Publication No.: WO 00/36127
(30) Priorité/Priority: 1998/12/17 (98/16163) FR

(51) Cl.Int.^{7}/Int.Cl.^{7} C12N 15/82, A01H 5/00, C12N 9/10, C12N 5/10

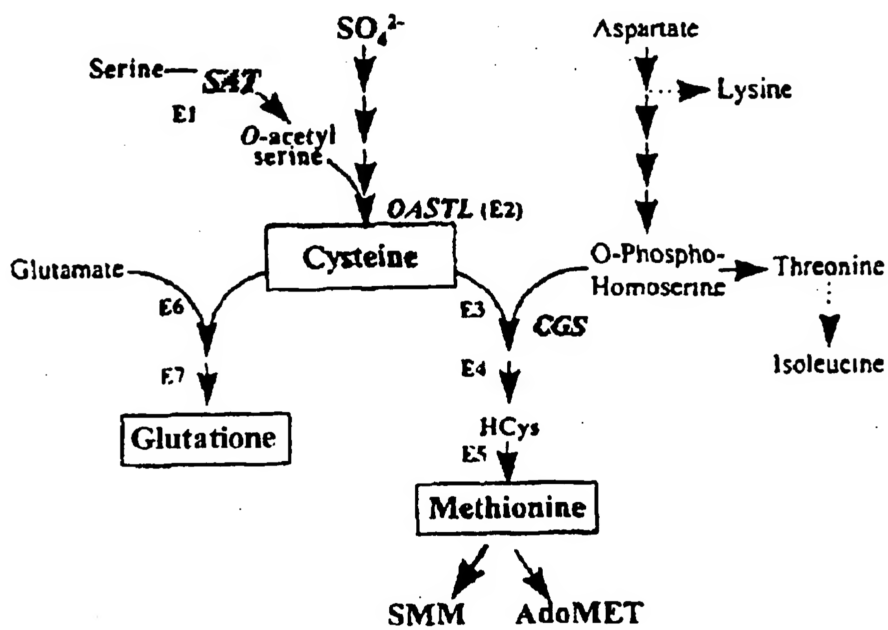
(71) Demandeur/Applicant:
AVENTIS CROPSCIENCE S.A., FR

(72) Inventeurs/Inventors:
JOB, DOMINIQUE, FR;
DEROSE, RICHARD, FR;
LAPPARTIENT, ANNE, FR;
DROUX, MICHEL, FR

(74) Agent: ROBIC

(54) Titre : PROCEDE POUR AUGMENTER LA TENEUR EN COMPOSES SOUFRES ET NOTAMMENT EN CYSTEINE, METHIONINE ET GLUTATHION CHEZ LES PLANTES ET PLANTES OBTENUES

(54) Title: METHOD FOR INCREASING THE CONTENT IN SULPHUR COMPOUNDS AND PARTICULARLY IN CYSTEINE, METHIONINE AND GLUTATHIONE IN PLANTS AND PLANTS OBTAINED



Sequence illustrating the synthetic pathway for cysteine and sulphur derivatives (glutathione and methionine)

(57) Abrégé/Abstract:
The invention concerns a method for increasing the production of cysteine, methionine, glutathione and their derivatives in plant cells and plants. Said method consists in over-expression of a serine acetyltransferase (SAT) in the plant cells. The invention also concerns plants containing said plant cells.

PCT ORGANISATION MONDIALE DE LA PROPRIETE INTELLECTUELLE
Bureau international

DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets [7] : C12N 15/82, 5/10, A01H 5/00, C12N 9/10	A1	(11) Numéro de publication internationale: **WO 00/36127** (43) Date de publication internationale: 22 juin 2000 (22.06.00)

(21) Numéro de la demande internationale: PCT/FR99/03179

(22) Date de dépôt international: 17 décembre 1999 (17.12.99)

(30) Données relatives à la priorité:
98/16163 17 décembre 1998 (17.12.98) FR

(71) Déposant *(pour tous les Etats désignés sauf US)*: AVENTIS CROPSCIENCE S.A. [FR/FR]; 55, avenue René Cassin, F-69009 Lyon (FR).

(72) Inventeurs; et
(75) Inventeurs/Déposants *(US seulement)*: DROUX, Michel [FR/FR]; 32, avenue de Lauterbourg, F-69169 Tassin la Demi Lune (FR). LAPPARTIENT, Anne [FR/FR]; 19A, rue Philippe Gonnard, F-69001 Lyon (FR). DEROSE, Richard [FR/FR]; 31, rue du Bois Guillaume, F-91000 Evry (FR). JOB, Dominique [FR/FR]; 181, rue Duguesclin, F-69003 Lyon (FR).

(74) Représentant commun: AVENTIS CROPSCIENCE S.A.; Boîte Postale 9163, F-69263 Lyon Cedex 9 (FR).

(81) Etats désignés: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, brevet ARIPO (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Publiée
Avec rapport de recherche internationale.

(54) Title: METHOD FOR INCREASING THE CONTENT IN SULPHUR COMPOUNDS AND PARTICULARLY IN CYSTEINE, METHIONINE AND GLUTATHIONE IN PLANTS AND PLANTS OBTAINED

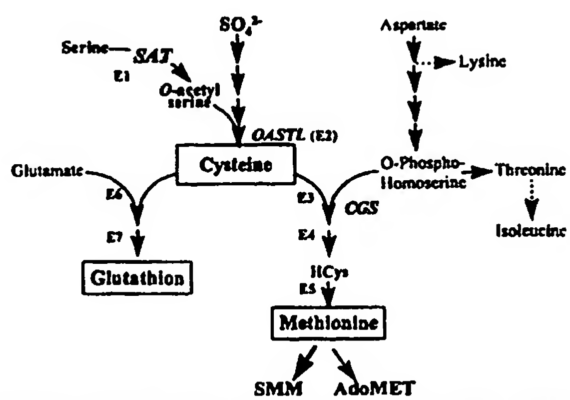
(54) Titre: PROCEDE POUR AUGMENTER LA TENEUR EN COMPOSES SOUFRES ET NOTAMMENT EN CYSTEINE, METHIONINE ET GLUTATHION CHEZ LES PLANTES ET PLANTES OBTENUES

(57) Abstract

The invention concerns a method for increasing the production of cysteine, methionine, glutathione and their derivatives in plant cells and plants. Said method consists in over-expression of a serine acetyltransferase (SAT) in the plant cells. The invention also concerns plants containing said plant cells.

(57) Abrégé

La présente invention concerne un procédé pour augmenter la production de cystéine, de méthionine, de glutathion et de leurs dérivés par les cellules végétales et les plantes, ledit procédé consistant à surexprimer une SAT dans les cellules végétales, et les plantes contenant lesdites cellules végétales.



Séquence illustrant la voie de synthèse de la cystéine et des dérivés soufrés (glutathion et méthionine).

SEQUENCE ILLUSTRATING THE SYNTHETIC PATHWAY OF CYSTEINE AND SULPHUR DERIVATIVES (GLUTATHION AND METHIONIN)

METHOD FOR INCREASING THE CONTENT IN SULPHUR COMPOUNDS AND PARTICULARLY IN CYSTEINE, METHIONINE AND GLUTATHIONE IN PLANTS AND PLANTS OBTAINED

Methionine is the first limiting essential amino acid in plants, in particular the leguminous plants which are one of the basic elements of the animal diet. Cysteine, another sulphur-containing amino acid, is not an essential amino acid, but can be taken to be a limiting element for animal nutrition since cysteine is derived, in animals, from methionine. In maize, the sulphur-containing amino acids are also limiting amino acids after lysine and tryptophan. The reason for this is that the major storage proteins of the seeds of these plants are lacking in these amino acids. The overproduction of methionine and cysteine in the seeds of leguminous plants (soybean, lucerne, pea, etc.) and of maize will thus have a considerable impact on the nutritional quality of these seeds.

So far, the increase in the nutritional quality of foods derived from the seeds of leguminous plants has been obtained by supplementation with chemically synthesized free methionine. For example, the average contents of methionine + cysteine in soybean and pea are of the order of 20 mg per g of protein. This content must be increased to a value of the order of 25 mg cysteine + methionine/g of protein to cover the dietary needs of a human adult, and to a

value of the order of 48 mg of cysteine + methionine/g of protein to cover those of pigs (De Lumen, B.O., Food Technology (1997) 51, 67-70).

The techniques for characterizing proteins enriched in sulphur-containing amino acids and the preparation of transgenic plants allowing the expression of such proteins, so as to increase the sulphur-containing amino acid content of these plants and thus their nutritive value for the animal diet, and thus to diminish the amount of synthesized methionine supplied, are now well known and described in the literature ([1] Korit, A.A. et al., Eur. J. Biochem (1991) 195, 329-334; WO 98/20133; WO 97/41239; WO 95/31554; WO 94/20828; WO 92/14822).

The enrichment in proteins with a high sulphur-containing amino acid content by such an approach remains, however, limited by the capacity of plant cells and of plants to produce the said sulphur-containing amino acids required for the synthesis of the protein. The reason for this is that plants overexpressing a protein rich in methionine and cysteine in their seed, such as for example lupins expressing 8S albumin, contain a level of free methionine and cysteine, and also of glutathione, which is lower than that of control plants ([2] Tabe, L. & Droux, M., 4th Workshop on Sulphur Metabolism, in press).

In the same way, peptides rich in sulphur-containing amino acids and having antifungal or antibacterial activity have been identified (WO 97/30082, WO 99/02717, WO 99/09184, WO 92/24594, WO 99/53053). The expression of these peptides in the plants makes it possible to increase the capacity of the said plants to resist certain fungal or bacterial attacks. Here again, the production of such peptides in the plants remains limited by the capacity of plant cells and plants to produce the sulphur-containing amino acids required for the synthesis of these peptides. The reason for this is that the expression of these peptides in the plant cell occurs to the detriment of the stock of glutathione, which is taken to be a reservoir for cysteine.

It has been observed that the limiting parameter of such an approach is indeed linked to this capacity to produce methionine or cysteine. It is therefore important to be able to modify in the plants this capacity to produce methionine and cysteine in sufficient quantities to allow the production of heterologous proteins with a high sulphur-containing amino acid content, that is to say to use a molecular strategy intended to increase the levels of cysteine and methionine in plants, and more particularly, crop plants of agronomical interest.

In plants, methionine biosynthesis is carried out from cysteine, this same cysteine being involved in the synthesis of glutathione.

Glutathione is a form of storage of reduced sulphur and represents 60 to 70% of the organic sulphur in the cell. Glutathione plays an important role for plants in the resistance to oxidative stress and in the elimination of toxic compounds. It thus participates in the elimination of xenobiotic compounds: heavy metals (for example) via the formation of phytochelatins and metallothionines; herbicides, via glutathione S-transferase activity; which are toxic to the plant, and in the plant's defence mechanisms against micro-organisms. By increasing a plant's cysteine content, and consequently its glutathione content, it is thus possible to modulate the plant's response to the different stresses mentioned above.

There are therefore two distinct metabolic pathways starting from cysteine, one for the preparation of methionine, the other for the preparation of glutathione (**Figure** 1) and for which the different enzymes involved are recalled below. The SAT (E1) and OASTL (E2) activities are at a metabolic crossroads between the assimilation of organic nitrogen and carbon (serine) and of inorganic sulphur (reduced sulphur from the sequence of assimilation and reduction of sulphate, shaded box). The cysteine is then incorporated into proteins, but also participates in

the synthesis of glutathione and methionine. The synthesis of the carbon backbone (*O*-phosphohomoserine) of this latter amino acid, is derived from aspartate. Aspartate is also the precursor for lysine, threonine and isoleucine synthesis. Moreover, the presence of a potentially limiting step for the synthesis of methionine by transcriptional regulation of CGS (cystathionine γ-synthase) is indicated in the diagram ([3] Giovanelli J. in Sulphur Nutrition and Sulphur Assimilation in Higher Plants, (1990) pp. 33-48; [4] Chiba Y. et al. (1999), Science, 286, 1371-1374). Methionine is the precursor of SAM (*S*-adenosylmethionine) which is involved in most methylation reactions, and of SMM (*S*-methylmethionine) taken to be a transport form and a storage form of methionine ([3]).

In plants the final steps of cysteine synthesis involve the two enzymes below:

E1) Serine acetyltransferase (EC 2.3.1.30)(SAT):
Serine + acetyl-coenzyme A → *O*-acetylserine + coenzyme A

E2) *O*-acetylserine (thiol) lyase (EC 4.2.99.8)(OASTL):
O-acetylserine + sulphide → cysteine + acetate

The synthesis of methionine from cysteine involves, successively, the three enzymes below:

E3) cystathionine γ-synthase (EC 4.2.99.9)(CGS):
O-phosphohomoserine + cysteine → cystathionine + Pi

Pi signifies inorganic phosphate.

E4) cystathionine β-lyase (EC 4.4.1.8)(CBL):

cystathionine + H_2O → homocysteine + pyruvate + NH_4^+

E5) methionine synthase (EC 2.1.1.14)(Ms):

homocysteine + 5-methyltetrahydrofolate → methionine + tetrahydrofolate

As for the synthesis of glutathione from cysteine, it involves, successively, the two enzymes below:

E6) γ-glutamylcysteine synthetase (EC 6.3.2.2)

glutamate + L-cysteine + ATP → γ-glutamylcysteine + ADP + Pi

E7) glutathione synthetase (EC 6.3.2.3)

γ-glutamylcysteine + glycine + ATP → glutathione + ADP + Pi

All these enzymes have been characterized and cloned in plants ([5] Lunn, J.E. et al., Plant Physiol. (1990) 94, 1345-1352; [6] Rolland, N. et al., Plant Physiol. (1992) 98, 927-935; [7] Droux, M. et al., Arch. Biochem. Biophys. (1992) 295, 379-390; [8] Rolland, N. et al., *Arch. Biochem* (1993) 300, 213-222; [9] Ruffet, M.L. et al., Plant Physiol. (1994) 104, 597-604; [10] Ravanel, S. et al., Arch. Biochem. Biophys. (1995) 316, 572-5584; [11] Droux, M. et al., Arch. Biochem. Biophys. (1995) 31, 585-595; [12] Ruffet, M.L. et al., Eur. J. Biochem. (1995) 227, 500-509; [13] Ravanel, S. et al., Biochem. J. (1996) 320, 383-392; [14] Ravanel, S. et al., Plant Mol. Biol. (1996) 29, 875-882; [15] Rolland, N. et al., Eur. J.

Biochem. (1996) 236, 272-282; [16] Ravanel, S. et al., Biochem. J. (1998) 331, 639-648; [17] Droux, M. et al., Eur. J. Biochem. (1998) 255, 235-245; [18] May, M.J., Leaver, C.J., Proc. Natl. Acad. Sci. USA (1994) 91, 10059-10063; [19] Ullmann, P. et al., Eur. J. Biochem. (1996) 236, 662-669; [20] Eichel, J. et al., Eur. J. Biochem. (1995) 230, 1053-1058).

It is known that for cysteine synthesis, the E1 and E2 enzymes are present in the three compartments of the plant cell, that is to say, the plasts, the cytosol and the mitochondria (5-6, 9, 12). These three E1 enzymes are named SAT2 and SAT4 for the (putative) chloroplast enzyme, and SAT1 for the mitochondrial enzyme, and SAT3 and SAT3' (SAT52) for the cytoplasmic enzyme. These localization attributions are based on sequence analysis.

For the methionine synthesis enzymes, the situation is different since the E3 and E4 enzymes are exclusively localized in the plasts (10-11, 13-14, 16), while the terminal E5 enzyme is in the cytosol (20).

As for the enzymes associated with the glutathione biosynthetic pathway, they are localized both in the chloroplast and in the cytosol ([21] Hell, R. and Bergmann, L., Planta (1990) 180, 603-612).

The E3 enzyme, of the methionine synthetic pathway, has a K_m (substrate concentration giving the half-maximal rate) of the order of 200 μM to 500 μM for

cysteine (10, 16, [22] Kreft, B-D. et al., Plant Physiol. (1994) 104, 1215-1220).

The E6 enzyme, of the glutathione synthetic pathway, also has a high K_m for cysteine, of the order of 200 μM [21].

It has now been observed the chloroplast serine acetyltransferase enzyme (**Figure** 2) and to a lesser degree the mitochondrial SAT are inhibited by cysteine, in contrast to the cytoplasmic enzyme (**Figure** 2), this inhibition constituting the essential limiting factor in the synthesis of cysteine in plant cells and being downstream of the methionine and glutathione.

The present invention thus consists in increasing the level of cysteine and methionine synthesized in the cellular compartments of plant cells, and in particular in the chloroplast compartment. Increasing the level of cysteine, the sulphur-containing precursor of glutathione and of methionine and its derivatives, advantageously makes it possible to increase the level of methionine and/or of glutathione in the plant cells and plants, and subsequently to improve the production of proteins, natural or heterologous, enriched in sulphur-containing amino acids in the plant cells and plants, and similarly the tolerance of the plants to different forms of glutathione-regulated stress.

This increase according to the invention is obtained by overexpressing a serine acetyltransferase (SAT) in the plant cells and plants.

The present invention thus relates to a method for increasing the production of cysteine, glutathione, methionine and sulphur-containing derivatives thereof, by plant cells and plants, the said method consisting in overexpressing an SAT in the plant cells and in plants containing the said plant cells.

The overexpressed SAT can consist of any SAT, whether of plant origin, in particular SAT2 or SAT4, SAT1, SAT3, SAT3' (SAT52), or of any other origin, in particular bacterial, in a native or mutant form or deleted of a fragment, and functional in the synthesis of *O*-acetylserine.

In particular, it can be a cysteine-sensitive SAT, such as for example a plant SAT, for example a chloroplast or mitochondrial SAT (SAT2, SAT4, SAT1), or a native SAT of bacterial origin ([22] Nakamori et al., 1998, Appl. Environ, Microbiol., 64, 1607-1611; [23] Takagi H. et al., 1999, Febs Lett. 452, 323-327; [24] Mino K. et al., 1999, Biosci. Biotechnol. Biochem., 63, 168-179).

It can also be a cysteine-insensitive SAT, such as, for example, a plant SAT, for example a cytoplasmic SAT (SAT3), or a mutant SAT of bacterial origin, made insensitive to cysteine by mutagenesis

([22] and [23], whose contents are incorporated here by reference), or any mutant plant SAT which is functional in the synthesis of *O*-acetylserine (the carbon-containing precursor for cysteine synthesis).

According to a specific embodiment of the invention, the SAT is an *Arabidopsis thaliana* SAT [12].

According to a first embodiment of the invention, the SAT is overexpressed in the cytoplasm of the plant cells. The SAT is either a plant cytoplasmic SAT, in particular the SAT3 (L34076) or SAT3′ or SAT52 (U30298), represented by the SEQ ID NO 1 or the SEQ ID NO 2, respectively, or an SAT of bacterial origin as defined above. The SAT overexpressed in the cytoplasm can also be a noncytoplasmic plant SAT, for example a chloroplast or mitochondrial SAT. These noncytoplasmic plant SATs, naturally, are expressed in the cytoplasm in the form of a precursor protein comprising a signal for addressing to the cellular compartment, other than the cytoplasm, into which the mature functional SAT is released. In order to overexpress these mature functional SATs in the cytoplasm, their addressing signal is removed. In this case, the SAT protein overexpressed in the cytoplasm is a noncytoplasmic plant SAT, with its signal(s) for addressing to cellular compartments, other than the cytoplasm, removed.

According to a specific embodiment of the invention, the noncytoplasmic SAT with its addressing signal removed is SAT1' represented by SEQ ID NO 3.

According to a second embodiment of the invention, the SAT is overexpressed in the mitochondria. The protein is advantageously expressed in the cytoplasm in the form of a signal peptide/SAT fusion protein, the mature functional SAT being released inside the mitochondria. Advantageously, the mitochondrial addressing signal peptide is made up of at least one mitochondrial addressing signal peptide from a plant protein which is located in mitochondria, such as the tobacco ATPase β-F1 subunit signal peptide [[25] Hemon P. et al. 1990, Plant Mol. Biol. 15, 895-904], or the SAT1 signal peptide represented by amino acids 1 to 63 in SEQ ID NO 4.

According to a specific embodiment of the invention, the mitochondrial SAT is SAT1 (U22964) represented by SEQ ID NO 4.

According to a third embodiment of the invention, the SAT is overexpressed in the chloroplasts of the plant cells.

The SAT will be expressed in the chloroplasts by any appropriate means, in particular by any means known to persons skilled in the art and widely described in the prior art.

According to a specific embodiment of the invention, the SAT is overexpressed in the chloroplasts

by integrating into the chloroplast DNA a chimeric gene comprising a DNA sequence encoding the said SAT, under the control of 5′ and 3′ regulatory elements that function in the chloroplasts. The techniques for insertion of genes into chloroplasts, such as the regulatory elements appropriate for the expression of the said genes in chloroplasts, are well known to persons skilled in the art and in particular are described in the following patents and patent applications: US 5,693,507, US 5,451,513 and WO 97/32977.

According to another specific embodiment of the invention, the SAT is overexpressed in the cytoplasm in the form of a transit peptide/SAT fusion protein, the function of the transit peptide being to address the SAT to which it is fused to the chloroplasts, the mature functional SAT being released inside the chloroplasts after cleavage at the chloroplast membrane.

In this case, the SAT can be a chloroplast SAT of plant origin, such as SAT2 or SAT4, represented by SEQ ID NO 5 or 6, respectively.

The SAT can also be a cytoplasmic SAT of plant origin or an SAT of bacterial origin as defined above. The cytoplasmic SATs are understood to include also noncytoplasmic SATs from which have been removed their signal for addressing to a compartment other than the cytoplasm, as defined above.

The transit peptides, their structures, their methods of functioning and their use in the construction of chimeric genes for addressing a heterologous protein into chloroplasts, as well as chimeric transit peptides comprising several transit peptides, are well known to persons skilled in the art and widely described in the literature. In particular, the following patent applications are mentioned: EP 189 707, EP 218 571 and EP 508 909, and the references cited in these patent applications, whose contents are incorporated here by reference.

In the fusion protein according to the invention, the SAT can be homologous or heterologous to the transit peptide. In the first case, the fusion protein is the SAT2 or the SAT4 protein expressed naturally in the chloroplasts of plant cells. In the second case, the transit peptide can be a transit peptide from an SAT2, represented by amino acids 1 to 32 of SEQ ID 5, or the transit peptide from an SAT4, represented by amino acids 1 to 30 of SEQ ID NO 6, or alternatively a transit peptide from another protein, which is located in plastids, in particular the transit peptides defined below. Protein which is located in plastids is understood to mean a protein expressed in the cytoplasm of plant cells in the form of a transit peptide/protein fusion protein, the mature protein being localized in the chloroplast after cleavage of the transit peptide.

A plant EPSPS transit peptide is, in particular, described in Patent Application EP 218,571, while a plant RuBisCO ssu transit peptide is described in Patent Application EP 189,707.

According to another embodiment of the invention, the transit peptide also comprises, between the C-terminal region of the transit peptide and the N-terminal region of the SAT a portion of sequence from the mature N-terminal region of a protein which is located in plastids, this portion of sequence generally comprising less than 40 amino acids from the N-terminal region of the mature protein, preferably less than 30 amino acids, more preferably between 15 and 25 amino acids. Such a transit peptide comprising a transit peptide fused to a part of the N-terminal region of a protein which is located in plastids is, in particular, described in Patent Application EP 189,707, more particularly for the transit peptide and the N-terminal region of plant RuBisCO ssu.

According to another embodiment of the invention, the transit peptide also comprises, between the C-terminal region of the N-terminal region of the mature protein and the N-terminal region of the SAT, a second transit peptide from a plant protein which is located in plastids. Preferably, this chimeric transit peptide comprising a combination of several transit peptides, is an optimized transit peptide (OTP) made by fusing a first transit peptide with a portion of

sequence from the mature N-terminal region of a protein which is located in plastids, which is fused with a second transit peptide. Such an optimized transit peptide is described in Patent Application EP 508,909, more particularly, the optimized transit peptide comprising the sunflower RuBisCO ssu transit peptide fused to a peptide made of the 22 N-terminal amino acids of the mature maize RuBisCO ssu, fused to the maize RuBisCO ssu transit peptide.

The present invention also relates to a transit peptide/SAT fusion protein in which the SAT defined above is heterologous to the transit peptide and in which the transit peptide is made of at least one transit peptide from a natural plant protein which is located in plastids, as defined above.

The present invention also relates to a nucleic acid sequence encoding a transit peptide/SAT fusion protein, described above. According to the present invention, "nucleic acid sequence" is understood to mean a nucleotide sequence which can be of DNA or RNA type, preferably of DNA type, in particular double-stranded, whether of natural or synthetic origin, in particular a DNA sequence in which the codons encoding the fusion protein according to the invention have been optimized according to the host organism in which it will be expressed, these optimization methods being well known to persons skilled in the art.

A subject of the invention is also the use of a nucleic acid sequence encoding an SAT according to the invention defined above, in particular for chloroplast, mitochondrial or cytoplasmic addressing, in a method for transforming plants, as a coding sequence allowing the modification of the cysteine, methionine, methionine derivatives, and glutathione contents of the transformed plants. This sequence can of course also be used in combination with other marker gene(s) and/or coding sequence(s) for one or more other agronomic properties.

The present invention also relates to a chimeric gene (or expression cassette) comprising a coding sequence as well as heterologous 5′ and 3′ regulatory elements capable of functioning in a host organism, in particular plant cells or plants, the coding sequence comprising at least one nucleic acid sequence encoding an SAT as defined above.

Host organism is understood to mean any monocellular or pluricellular higher or lower organism, into which the chimeric gene according to the invention can be introduced. They are in particular bacteria, for example *E. coli*, yeasts, in particular of the genera *Saccharomyces*, *Kluyveromyces* or *Pichia*, fungi, in particular *Aspergillus*, a baculovirus, or preferably plant cells and plants.

"Plant cell" is understood to mean according to the invention any cell derived from a plant and

capable of constituting undifferentiated tissues such as calli, differentiated tissues such as embryos, plant portions, plants or seeds.

"Plant" is understood to mean according to the invention any differentiated multicellular organism capable of photosynthesis, in particular monocotyledonous or dicotyledonous plants, more particularly crop plants intended or not as animal feed or for human consumption, such as maize, wheat, rape, soybean, rice, sugar cane, beet, tobacco, cotton and the like.

The regulatory elements required for the expression of the a nucleic acid sequence encoding a fusion protein according to the invention are well known to persons skilled in the art according to the host organism. They comprise, in particular, promoter sequences, transcription activators, termination sequences including start and stop codons. The means and methods of identifying and selecting the regulatory elements are well known to persons skilled in the art and widely described in the literature.

The invention relates more particularly to the transformation of plants. Promoter regulatory sequences which can be used in plants, are any promoter sequence of a gene which is naturally expressed in plants, in particular a promoter which is expressed in particular in the leaves of plants such as, for example, so-called constitutive promoters of bacterial,

viral or plant origin, or alternatively so-called light-dependent promoters such as that of a plant ribulose-biscarboxylase/oxygenase (RuBisCO) small subunit gene or any suitable known promoter that can be used. Among promoters of plant origin which can be mentioned are the histone promoters as described in Application EP 0,507,698, or the rice actin promoter (US 5,641,876). Among promoters of plant virus genes which can be mentioned are that of the cauliflower mosaic (CAMV 19S or 35S), or the circovirus promoter (AU 689 311).

It is also possible to use a promoter regulatory sequence which is specific for regions or tissues specific to plants, and more particularly seed-specific promoters ([26] Datla, R. et al., Biotechnology Ann. Rev. (1997) 3, 269-296), in particular the napin (EP 255,378), phaseolin, glutenin, zein, helianthinin (WO 92/17580), albumin (WO 98/45460), oelosin (WO 98/45461), ATS1 or ATS3 (WO 99/20275) promoters.

According to the invention, it is also possible to use, in combination with the regulatory promoter sequence, other regulatory sequences which are situated between the promoter and the coding sequence, such as transcription enhancers, such as, for example the translational enhancer of tobacco mosaic virus (TMV) described in Application WO 87/07644, or of

tobacco etch virus (TEV) described by Carrington & Freed.

Regulatory termination or polyadenylation sequences which can be used, are any corresponding sequence of bacterial origin, such as for example the nos terminator of *Agrobacterium tumefaciens*, or alternatively of plant origin, such as for example a histone terminator as described in Application EP 0,633,317.

The present invention also relates to a cloning and/or expression vector for the transformation of a host organism containing at least one chimeric gene as defined above. This vector comprises, besides the chimeric gene above, at least one origin of replication. This vector can be a plasmid, a cosmid, a bacteriophage or a virus, which has been transformed by introducing a chimeric according to the invention. Such transformation vectors, according to the host organism to be transformed, are well known to persons skilled in the art and widely described in the literature. For the transformation of plant cells or plants, a virus, moreover containing its own elements of replication and expression, can, in particular, be used to transform developed plants. Preferably, the transformation vector of plant cells or plants according to the invention is a plasmid.

For the transformation of host organisms, the chimeric gene according to the invention can be used in

combination with a selection marker gene, either in the same vector, the two genes being combined in a convergent, divergent or colinear manner, or alternatively in two vectors used simultaneously for transforming the host organism. Such marker genes and their use for transforming host organisms are well known to persons skilled in the art and widely described in the literature.

Among genes encoding selection markers which can be mentioned are antibiotic-resistance genes, genes which impart tolerance to herbicides (bialaphos, glyphosate or isoxazoles), genes encoding easily identifiable enzymes such as the GUS enzyme (or GFP, "Green Fluorescent Protein"), or genes encoding pigments or enzymes which regulate the production of pigments in the transformed cells. Such selection marker genes are in particular described in Patent Applications EP 242 236, EP 242 246, GB 2 197 653, WO 91/02071, WO 95/06128, WO 96/38567 or WO 97/04103.

The subject of the invention is also a method for transforming host organisms, in particular plant cells, by integration of at least one nucleic acid sequence or one chimeric gene as defined above, which transformation may be obtained by any known appropriate means, widely described in the specialist literature and in particular the references cited in the present application, more particularly by the vector according to the invention.

One series of methods consists in bombarding cells, protoplasts or tissues with particles to which the DNA sequences are attached. Another series of methods consists in using, as a means of transferring into the plant, a chimeric gene inserted into an *Agrobacterium tumefaciens* Ti plasmid or an *Agrobacterium rhizogenes* Ri plasmid. Other methods can be used, such as microinjection or electroporation, or alternatively direct or PEG precipitation. Persons skilled in the art will choose the appropriate method according to the nature of the host organism, in particular of the plant cell or of the plant.

The subject of the present invention is also the host organisms, in particular plant cells or plants, which are transformed and which contain a chimeric gene defined above.

The subject of the present invention is also the plants containing transformed cells, in particular the plants regenerated from the transformed cells. The regeneration is obtained by any appropriate method which depends on the nature of the species, as for example described in the above references. Patents and patent applications which are mentioned for the methods of transforming plant cells and of regenerating plants are, in particular, the following: US 4,459,355, US 4,536,475, US 5,464,763, US 5,177,010, US 5,187,073, EP 267,159, EP 604,662, EP 672,752, US 4,945,050, US 5,036,006, US 5,100,792, US 5,371,014, US 5,478,744,

US 5, 179,022, US 5,565,346, US 5,484,956, US 5,508,468, US 5,538,877, US 5,554,798, US 5,489,520, US 5,510,318, US 5,204,253, US 5,405,765, EP 442,174, EP 486,233, EP 486,234, EP 539,563, EP 674,725, WO 91/02071 and WO 95/06128.

The subject of the present invention is also the transformed plants derived from the cultivation and/or the crossing of the above regenerated plants, as well as the seeds of the transformed plants.

The transformed plants which can be obtained according to the invention can be of monocotyledonous type, such as for example cereals, sugar cane, rice and maize, or of dicotyledonous type, such as for example tobacco, soybean, rape, cotton, beet, clover, etc.

The plants transformed according to the invention can contain other genes of interest, which confer novel agronomic properties on the plants. Among genes conferring novel agronomic properties on the transformed plants which can be mentioned are genes conferring tolerance to certain herbicides, those conferring tolerance to certain insects, and those conferring tolerance to certain diseases. Such genes are in particular described in Patent Applications WO 91/02071 and WO 95/06128. Mention may also be made of genes which modify the composition of the modified plants, in particular the content and quality of certain essential fatty acids (EP 666,918), or alternatively the content and quality of proteins, in

particular in the leaves and/or seeds of the said plants. In particular, genes encoding proteins enriched in sulphur-containing amino acids are cited([1]; WO 98/20133; WO 97/41239; WO 95/31554; WO 94/20828; WO 92/14822; US 5,939,599, US 5,912,424). The function of these proteins enriched in sulphur-containing amino acids is also to trap and store surplus cysteine and/or methionine, making it possible to avoid any problems of toxicity linked to an overproduction of these sulphur-containing amino acids, by trapping them.

Mention may also be made of genes encoding peptides rich in sulphur-containing amino acids and more particularly rich in cysteine, the said peptides also having antibacterial and/or antifungal activity. More particularly, plant defensins are mentioned, as well as lytic peptides of any origin, and more particularly the following lytic peptides: androctonin (WO 97/30082 and WO 99/09189), drosamicin (WO 99/02717), thanatin (WO 99/24594) or heliomicin (WO 99/53053).

These other genes of interest can be combined with the chimeric gene according to the invention either by conventional crossing of two plants each containing one of the genes (the first being the chimeric gene according to the invention and the second being the gene encoding the protein of interest), or by transforming the plant cells of a plant containing the

gene encoding the protein of interest, with the chimeric gene according to the invention.

The following examples illustrate the invention, without, however, looking to limit its scope.

All of the methods or operations described below in these examples are given by way of examples and correspond to a choice made from the different methods available to arrive at the same result. This choice has no bearing on the quality of the result and consequently, any adapted method can be used by persons skilled in the art to arrive at the same result. Most of the methods for engineering DNA fragments are described in "Current Protocols in Molecular Biology" Volumes 1 and 2, Ausubel F.M. et al, published by Greene Publishing Associates and Wiley Interscience (1989) or in Molecular Cloning, T. Maniatis, E.F. Fritsch, J. Sambrook, 1982.

The contents of all the references cited in the above description and in the following examples are incorporated into the text of the present patent application by reference.

Example 1. Demonstration of the inhibition of chloroplast serine acetyltransferase from pea (*Pisum sativum*) leaves by cysteine

The three subcellular compartments corresponding to the cytosol (preparation from a

subcellular fractionation of pea protoplasts, [12]), to mitochondria and to chloroplasts are prepared from pea leaves [12]. The soluble proteins are extracted therefrom and the serine acetyltransferase activity present in each of the compartments is measured by means of a described technique [12, 17].

Description of the assay method:

The serine acetyltransferase activity is measured by high performance liquid chromatography (HPLC), by assaying the *O*-acetylserine formed during the course of the reaction (reaction 1), after derivatization with orthophthalaldehyde (OPA). A defined quantity of the protein extract, corresponding to the cytosol and to the different soluble fractions of chloroplasts (stroma) and of mitochondria (matrix), is desalted on a PD10 column (Pharmacia) pre-equilibrated in a buffer containing 50 mM Hepes-HCl, pH 7.5 and 1 mM EDTA. The enzyme reaction is carried out in the presence of 50 mM Hepes-HCl, pH 7.5, 1 mM dithiothreitol, 10 mM L-serine, 2.5 mM acetyl-CoA, in a 100 μl reaction volume, at 25°C. After 10 to 15 minutes' incubation, the reaction is stopped by addition of 50 μl of 20% (W/V) trichloroacetic acid. The proteins thus precipitated are then eliminated by centrifugation for 2 min at 15,000 g. The supernatant, which contains the reaction product (OAS), is mixed with 500 μl of a derivatization solution (54 mg of OPA

dissolved in 1 ml of absolute ethanol, 9 ml of a 400 mM solution of borate-NaOH, pH 9.5, and 0.2 ml of 14 M β-mercaptoethanol) and incubated for 2 min. A fraction of this mixture (20 μl) is injected onto a reverse phase column (3.9 × 150 mm, AccQ Tag C_{18} column, Waters) which is connected to an HPLC system. The buffers used to elute the compounds derivatized by OPA are: Buffer A, 85 mM sodium acetate, pH 4.5 and 6% (V/V) acetonitrile, pH 4.5; Buffer B, 60% (V/V) acetonitrile in water. The *O*-acetylserine, which has been derived by OPA, is eluted with a continuous linear gradient of buffer B in buffer A, of 25 to 70% (V/V), and is detected by fluorescence at 455 nm (excitation at 340 nm). The retention time of *O*-acetylserine, under our conditions, is of the order of 6.2 min., and the amount of product which is formed in the enzyme assays is quantified from a standard curve which is obtained for *O*-acetylserine. The enzyme assays were optimized in order to respect the optimum pH of the reaction, the linearity with time, and in order to operate under saturating conditions of substrates.

<u>Effect of cysteine on serine acetyltransferase activity of pea leaves</u>

Incubation (2 min) is carried out in the presence of protein extract (cytosol, matrix, and stroma), and in the presence of increasing concentrations of L-cysteine (from 0 to 1 mM), before

adding saturating concentrations of the serine acetyltransferase substrates, L-serine (10 mM) and acetyl-CoA (2.5 mM). The enzyme reaction and assay of residual *O*-acetylserine in the supernatant are carried out as described above. The result of these experiments is represented in the graph of **Figure** 2, in the annex.

If free cysteine (from 0 to 1 mM, **Figure** 2) is added to the different assays, a very strong inhibition of chloroplast serine acetyltransferase activity is observed (inhibition constant of the order of 30 μM). Mitochondrial serine acetyltransferase activity is inhibited, but at higher concentrations of cysteine (inhibition constant of the order of 300 μM). On the other hand, cytosolic serine acetyltransferase activity is insensitive to inhibition by cysteine up to concentrations of the order of 1 mM cysteine (**Figure** 2). This result proves, therefore, that only chloroplast serine acetyltransferase activity, thus the enzyme associated with the sulphate assimilation pathway, is inhibited by the final product, L-cysteine.

Table I: Determination of the specific activities and IC_{50} values of cysteine for each of the serine acetyltransferase isoforms.

Serine acetyltransferase (*Pisum sativum*)		
	Specific activity nmol OAS·min^{-1}·mg^{-1}	IC_{50} L-cysteine μM
Stroma	0.93 ± 0.2	33.4 ± 8
Matrix	10 ± 2	283 ± 50
Cytosol	0.83 ± 0.3	no inhibition

The concentration of L-cysteine which makes it possible to obtain 50% inhibition (IC_{50}) under standard reaction conditions, and which is calculated for different enzyme preparations, is represented in Table I. Determination of the serine acetyltransferase enzyme activity and of the IC_{50} is carried out for 9 different experiments (on stroma), and for 3 experiments for the cytosolic extracts and 3 for the mitochondrial extracts. Similarly, activity of chloroplast serine acetyltransferase from spinach leaves is cysteine-sensitive. Conversely, in *Arabidopsis thaliana*, only the activity of the isoform associated with the cytosolic compartment seems to be controlled by the level of cysteine ([27] Noji M. et al. 1998, J. Biol. Chem. 273, 32739-32745; [28] Inoue K. et al. 1999, Eur. J. Biochem. 266, 220-227). For

these authors, the activity associated with the chloroplast compartment is cysteine-insensitive. It would seem, therefore, that inhibition of the chloroplast serine acetyltransferase activity by cysteine is a plant-specific phenomenon, but, in particular, is very pronounced in leguminous plants, such as pea.

Study of the mode of inhibition of serine acetyltransferase activity by cysteine

The enzyme reaction rate was determined for fixed concentrations of cysteine (0 μM; 10 μM; 20 μM; 40 μMm 60 μM and 100 μM), by varying either the L-serine concentration or the acetyl-CoA concentration, for saturating concentrations of the second co-substrate. For each of the kinetics obtained, the affinity of the enzyme for these substrates does not seem to be affected, but, on the other hand, the maximum reaction rate is modified. The more the concentration of L-cysteine increases, the more the rate of *O*-acetylserine synthesis decreases. For each of the conditions analysed, the inhibition constant K_i was estimated to be of the order of 30 (±2.2) μM (variable substrate: serine), and 22 (±2) μM (variable substrate: acetyl-CoA). We were able to show that cysteine is a non-competitive type of inhibitor of serine acetyltransferase activity and that, moreover, it is an allosteric type inhibitor (Hill constant of the order

of 1.6±0.3 μM), using conventional kinetics equations ([29] Segel, I.H. (1995), John Wiley and Sons, New York). These results indicate that inhibition of the chloroplast enzyme takes place at a site other than the active site, which moreover, does not exist in the serine acetyltransferase isoform which is present in the cytosol.

These inhibition constants are consistent with the cysteine concentration determined for pea chloroplasts of 40 ± 10 μM (2 nmol/mg chlorophyll), a value which is calculated for a stromal compartment volume of the order of 35 to 65 μl per mg of chlorophyll.

Dissociation of the bi-enzymatic complex, cysteine synthase, by cysteine

The serine acetyltransferase of the plant cell, like its bacterial homologue, forms an enzymatic complex with *O*-acetylserine (thiol) lyase, the enzyme which catalyses the condensation of reduced sulphur with *O*-acetylserine. This bi-enzymatic complex is called cysteine synthase. All of the serine acetyltransferase of the chloroplast exists in a form complexed with *O*-acetylserine (thiol) lyase, while the majority of the *O*-acetylserine (thiol) lyase is in the free form. The distribution of each of these enzymes in each of the subcellular compartments of pea leaves is described in Table II.

Table II: Specific activity of serine acetyltransferase and *O*-acetylserine (thiol) lyase activities in the cellular compartments of pea leaves

	Serine aceytl-transferase	*O*-acetylserine (thiol) lyase	
	Specific activity (mU/mg)		OASTL/SAT Ratio
Stroma	0.85	260	306
Matrix	12	50	4
Cytosol	0.90	180	200

The ratio of *O*-acetylserine (thiol) lyase (OASTL) activity to serine acetyltransferase (SAT) activity reflects the large excess of OASTL over SAT. In particular in the stoma (chloroplast), where the assimilation and reduction of sulphate takes place, and in the cytosol, 95% of the OASTL activity is in the free form. These conditions are necessary for optimal synthesis of cysteine [14]. The cysteine synthase complex is composed of a serine acetyltransferase tetramer and two *O*-acetylserine (thiol) lyase dimers. *O*-Acetylserine, the reaction product of serine acetyltransferase, dissociates this bienzymatic complex, and sulphur tends to stabilize it [14]. These protein-protein interactions within the complex confer novel properties on each of the enzymes; in particular serine acetyltransferase acquires novel catalytic

properties compared to the free form. Moreover, complexed *O*-acetylserine (thiol) lyase is inactive in cysteine synthesis, and only the free form (in excess in the cell) catalyses cysteine synthesis [14].

A chloroplast (*Pisum sativum*) fraction, pre-incubated in the presence of an optimal concentration of cysteine (0.1 mM), which inhibits serine acetyltransferase (see Figure 2), then undergoes gel filtration chromatography which allows the separation of molecules according to their molecular mass. Under these conditions the cysteine synthase complex dissociates into serine acetyltransferase tetramers and *O*-acetylserine (thiol) lyase dimers. Chloroplast serine acetyltransferase in its free form is still sensitive to inhibition by cysteine. To refine this result and to confirm that inhibition of the enzyme is not dependent upon interaction with OASTL, a serine acetyltransferase was partially purified from pea chloroplasts, by ion exchange chromatography followed by molecular filtration chromatography carried out in the presence of *O*-acetylserine (1 mM), a condition which leads to dissociation of the complex.

The serine acetyltransferase fraction thus free of contamination by *O*-acetylserine (thiol) lyase is incubated in the presence of increasing concentrations of cysteine under the conditions described in Table I and **Figure** 2. The calculated IC_{50} is of the order of 15 +/- 3 micromolar and is

comparable to the value obtained above for the enzyme under chloroplast conditions (see Table I). This latter result makes it possible to establish a model to explain the inhibition of chloroplast serine acetyltransferase. In **Figure** 3, the tetrameric form of serine acetyltransferase (SAT) is depicted by the grey circles and the *O*-acetylserine (thiol) lyase (OASTL) dimer by the black circles. The functional cysteine synthase complex in the cell is depicted by the combination of the two molecular populations. In the presence of cysteine, the cysteine synthase complex binds cysteine, which modifies the protein-protein interactions within the cysteine synthase complex, and leads to dissociation into SAT tetramers and OASTL dimers. The SAT thus in its free form is therefore sensitive to cysteine, and it is known that this structure has a tendency to form aggregates (apart from the cysteine synthase complex) whose effect is to cause a loss of its activity [14].

Example 2. Isolation and characterization of a gene encoding a putative cytoplasmic serine acetyltransferase isoform (SAT3) [12]

In this example the procedure described on page 502 of Ruffet et al. [12], is taken up, in particular the chapters described under the headings "Bacterial strain and growth conditions" and "Isolation

of *A. thaliana* serine acetyltransferase cDNA clones by complementation in *E. coli*".

A gene encoding a putative cytosolic serine acetyltransferase (Z34888 or L34076) represented in **Figure** 4 (SEQ ID NO 1), was isolated by functional complementation of an *Escherichia coli* strain deficient in serine acetyltransferase activity. Analysis of the primary sequence showed the presence of strong similarity with the sequence of the bacterial enzyme (56% homology and 41% identity).

The following primers were used to amplify the nucleotide sequence and to clone it into the vector used for transforming tobacco plants:

Oligo 1 : 5'GAGAGA<u>GGAT</u> <u>CC</u>TCTTTCCA ATCATAAACC ATGGCAACAT GCATAGACAC ATGC 3'

Oligo 2 : 5'GGCTCACCAG ACTAATACAC TAAATTGTGT TTAC<u>CTCGAG</u> AGAGAG 3'

These primers make it possible to introduce a 5' *BamH1* restriction site (GGATCC) and a 3' *Sac1* restriction site (GAGCTC).

The N terminus of the amino acid sequence of the SAT3 isoform does not have the characteristics of organelle (mitochondrion or chloroplast) addressing peptides. This analysis leads to the assumption that this isoform is located in the cytosol [12]. The absence of an addressing peptide of chloroplast type in this isoform was confirmed in chloroplast import experiments ([29] Murillo et al. 1995, Cell. and Mol.

Biol. Research 41, 425-433). Conversely, a study using constructs which include a portion of the nucleotide sequence and a marker protein (Green Fluorescent Protein, GFP) showed the presence of the fusion product (5′-SAT3-GFP) in the chloroplast of transformed *A. thaliana* plants (vegetative stage of the plant) and also in the cytosol (at the floral stage)[27].

The SAT3 gene (L34076) contains no introns.

Example 3. Overexpression and purification of SAT3 in *Escherichia coli*

The defined protocol for overexpression of the enzyme in *E. coli* makes it possible to purify the enzyme in its free form or complexed with plant *O*-acetylserine (thiol) lyase, the cysteine synthase complex [14]. Using the purified proteins, the effect of cysteine on serine acetyltransferase activity was analysed by a spectrophotometric assay based on the consumption of acetyl-CoA during reaction 1, as a function of incubation time. This analysis is carried out in a medium (1 ml) containing 50 mM Hepes-HCl, pH 7.5, 2 mM L-serine and 0.2 mM acetyl-CoA. The reaction is followed by measuring the decrease in absorbance at 232 nm (molecular extinction coefficient of 4200 $M^{-1}cm^{-1}$)([30] Kredich, N.M. et al., J. Biol. Chem. (1969) 244, 2428-2439). We were able to show that this isoform (SAT3) in its free form or complexed with *O*-acetylserine (thiol) lyase, is cysteine-insensitive.

This result allows us to confirm that this cDNA (L34076, **Figure** 4) encodes a cytosolic serine acetyltransferase, since the amino acid composition of the N-terminus does not have the characteristics of transit peptides, and moreover, since this serine acetyltransferase is cysteine-insensitive. This latter result is similar to observations which have been obtained for the cytosolic serine acetyltransferase activity of pea leaves (**Figure** 2 and Table I).

Example 4. Isolation and characterization of a gene encoding a cytoplasmic serine acetyltransferase isoform (SAT3')(U30298)

The procedure of Example 3 is repeated, using oligonucleotides 3 and 4 below:

Oligo 3 : 5'GAGAGAGGAT CCTCTTATCG CCGCGTTAAT ATGCCACCGG CCGGAGAACTC C 3'

Oligo 4 : 5'GAGCCTTACC AGTCTAATGT AGTATATTTC AACCTCGAGA GAGAG 3'

A gene is isolated which encodes an acetyltransferase (U 30298), and is represented in **Figure** 5 (SEQ ID NO 2). Analysis of the primary sequence showed the presence of strong similarity with the sequence of the bacterial enzyme (51.6% homology and 42.6% identity). The N-terminal structure (absence of the conditions necessary for organelle addressing) indicates that this isoform is located in the cytosol. On the other hand, it is given as being cysteine-

sensitive [27]. This result differs from the data obtained from pea leaves (and from spinach leaves), in the sense that the cysteine regulation site seems to be confined to the cytosol in *A. thaliana* [27]. Moreover, it would seem that *A. thaliana* has at least two cytosolic isoforms: SAT3 (Example 3) and SAT3' (or U30298, Example 4). Unlike the SAT3 gene, the gene corresponding to **SAT3'** has an intron.

Example 5. Isolation and characterization of genes encoding a serine acetyltransferase isoform (SAT1')

The procedure described in Example 3 is repeated for the present example.

A gene encoding a serine acetyltransferase (L78443), which is represented in **Figure** 6 (SEQ ID NO 3), was isolated by functional complementation of an *Escherichia coli* strain deficient in serine acetyltransferase activity [12]. Analysis of the primary sequence shows strong similarity with the sequence of the bacterial enzyme (52.7% homology and 39% identity).

The following primers were used to amplify the nucleotide sequence and to clone it into the vector which is used for transforming tobacco plants (in bold characters in **Figure** 3):

Oligo 5 : 5'GAGAGAGGAT CCCCTCCTCC TCCTCCTCCT ATGGCTGCGT GCATCGACAC CTG 3'

Oligo 6 : 5'GCTCACCAGC CTAATACATT AAACTTTTTC AGCTCGAGAG AGAG 3'

These primers make it possible to introduce a 5′ *BamH1* restriction site (GGATCC) and a 3′ *Sac1* restriction site (GAGCTC).

A second gene is obtained which encodes a putative mitochondrial serine acetyltransferase (U22964), and is represented in **Figure** 7 (SEQ ID NO 4), by repeating the same procedure, using oligo 7 to replace oligo 5 as the 5′ primer.

Oligo 7°: 5'GAGAGAGGAT CCGGCCGAGA AAAAAAAAAA ATGTTGCCGG TCACAAGTCG CCG 3'

The N-terminus of the amino acid sequence of the SAT1 isoform has the characteristics of organelle (mitochondrion or chloroplast) addressing peptides. Localization in the mitochondrion was recently confirmed by constructing a fusion protein which includes the 5′ portion and "green fluorescent protein" (5′SAT1-GFP) and by transforming *Arabidopsis thaliana* plants [27]. Neither the SAT1′ gene (L78443) nor the SAT1 gene (U22964), like its homologue (SAT3), has introns.

Example 6. Overexpression and purification of SAT1 in *Escherichia coli*. Localization of this isoform in *A. thaliana*

The defined protocol for overexpression of the enzyme in *E. coli* makes it possible to purify the enzyme (in its transit peptide-lacking form, SAT L78443) in its free form or complexed with plant *O*-acetylserine (thiol) lyase, the cysteine synthase complex [14]. Using the purified proteins, the effect of cysteine on serine acetyltransferase activity was analysed by spectrophotometric assay, based on the consumption of acetyl-CoA during reaction 1, as a function of incubation time (see Example 3). Analysis was also carried out by HPLC assay of the reaction product (OAS) (see Example 1). We were able to show that this isoform (SAT1'), in its free form or complexed with *O*-acetylserine (thiol) lyase, is cysteine-insensitive. This latter result parallels the observations obtained for pea leaf mitochondrial serine acetyltransferase activity (**Figure** 2 and Table I), the latter being inhibited at non-physiological concentrations of cysteine.

Using a preparation of mitochondria obtained from pea leaves or from protoplasts from cell cultures, localization in the mitochondrion was confirmed for this isoform.

A mitochondrial fraction lacking in plastid and in cytosolic contaminants was obtained by using the

protocol defined for pea leaf mitochondria [12]. The molecular mass of the polypeptide as revealed by antibodies raised against the peptide [-TKTLHTRPLLEDLDR-] (see SAT1 amino acid sequence), is of the order of 34,000 daltons, a value which is in agreement with the mass of the protein as obtained using sequence analysis programs for predicting cleavage sites.

Example 7. Isolation and characterization of genes encoding a serine acetyltransferase isoform (SAT2)

The procedure described for Example 3 is repeated for the present example.

A gene which encodes a serine acetyltransferase (L78444), represented in **Figure** 8 (SEQ ID NO 5), was isolated by functional complementation of an *Escherichia coli* strain deficient in serine acetyltransferase activity [12]. Analysis of the primary sequence showed the presence of strong similarity with the sequence of the bacterial enzyme (49.5% homology and 35.4% identity).

The following primers were used to amplify the nucleotide sequence and to clone it into the vector which was used to transform tobacco plants (in bold characters in **Figure** 8):

Oligo 8 : **5'GAGAGAGGAT CCGACAAGTT GGCATAATTT ATGGTGGATC TATCTTCCT 3'**

Oligo 9 : **5'CCTGTGTGAT TGTCGTGTAG TACTCTAGAA ACTCGAGAGA GAG 3'**

These primers make it possible to introduce a 5' *BamH1* restriction site (GGATCC) and a 3' *Sac1* restriction site (GAGCTC).

Analysis of the N-terminal portion of the sequence shows the presence of characteristics for addressing of the protein to an organelle (mitochondrion or chloroplast). Unlike the other isoforms described above, the SAT2 gene is complex and has several introns. Comparing SAT2 sequences with its homologues from *A. thaliana*, from plants and from other organisms, leads to the assumption of a prokaryotic origin (**Figure** 10). Moreover, analysis of the N-terminal sequence using the chloroP program [http://www.cbs.dtu.dk/services/chlorP/], indicates a high probability of the presence of a chloroplast-type transit peptide.

Example 8. Isolation and characterization of genes encoding a serine acetyltransferase (SAT4) isoform

A gene which encodes a serine acetyltransferase (SAT4), represented in **Figure** 9 (SEQ ID NO 6), was isolated by functional complementation of an *Escherichia coli* strain deficient in serine acetyltransferase activity [12]. Analysis of the

primary sequence showed the presence of strong similarity with the sequence of the bacterial enzyme (44.5% homology and 32% identity).

The following primers were used to amplify the nucleotide sequence and to clone it into the vector which was used for transforming tobacco plants:

Oligo 10 :5'GAGAGAGGAT CCGACAAGTTGG CATAATTTAT GGCTTGTATA AACGGCGAGA ATCGTGATTT TTCTT 3'

Oligo 11: 5'TACCTCGTAC CACTCAGAAC TCTAGAAACT CGAGAGAGAG3'

These primers make it possible to introduce a 5' *BamH1* restriction site (GGATCC) and a 3' *Sac1* restriction site (GAGCTC).

Analysis of the N-terminal portion sequence shows the presence of characteristics for addressing of the protein to an organelle (mitochondrion or chloroplast). The SAT4 gene, like that of SAT2, is complex and has several introns. Comparing SAT4 sequences with its homologues from *A. thaliana*, from plants and from other organisms, leads to the assumption of a prokaryotic origin (**Figure** 10). Moreover, analysis of the N-terminal sequence using the chloroP program [http://www.cbs.dtu.dk/services/chlorP/], indicates a high probability of the presence of a chloroplast-type transit peptide. **Figure 10** represents the sequence comparison and was carried out using the Clustaw program (Vector NTI software). SAT2 and SAT4 are closer to the prokaryotic SATs than are SAT3, SAT1 and SAT52.

Moreover, the branch also comprises an SAT from red alga (AB00848), which has been identified as a cysteine-sensitive protein located in the chloroplast ([32] Toda et al., 1998, Biochim. Biophys. Acta 1403, 72-84). SAT4 is identified as being on chromosome 4 (Bac clone F8D20, access number AL031135).

Example 8. Constructs used for transforming tobacco plants of the small Havanna variety

Transgene expression in leaves

Transformation of tobacco plants is carried out through *Agrobacterium tumefaciens* EHA105, which contains the pBI121 vector (Clontech) (**Figures** 11 and 12).

SAT3 (or SAT1′ or any cysteine-insensitive SAT)

To obtain expression of the SAT3 (SEQ ID NO 1) of Example **2 in the chloroplast (Figure** 11), an extension which allows addressing to this compartment is introduced 5′ of the cDNA. For this, the optimized transit peptide previously described is used.

A kanamycin-resistance gene (NPTII) which encodes neomycin phosphotransferase, and which is used as a selection marker for transforming tobacco, is cloned between the left (LE) and right (RE) edges of the t-DNA. Expression of the NPTII gene is under the control of the promoter and of the terminator of *A. tumefaciens* nopalin synthase. Downstream, the

β-glucuronidase gene which has been cloned between the unique *Bam*H1 and the unique *Sac*1 sites, is under the control of the cauliflower mosaic virus (CaMV) 35S promoter and the nopalin synthase gene polyadenylation signal from the Ti plasmid. The OTP-SAT3 construct is inserted between the *Xho* and *Sac*1 sites of the vector, from which has been deleted the β-glucuronidase gene (**Figure** 11).

<u>SAT1, SAT3, SAT3', SAT2, SAT4 or any SAT</u>

To obtain SAT expression in any of the subcellular compartments (cytosol, mitochondrion or chloroplast), the transgene is introduced into the appropriate vector, which is described in **Figure** 12.

A kanamycin-resistance gene (NPTII) which encodes neomycin phosphotransferase, and which is used as a selection marker for transforming tobacco, is cloned between the left (LE) and right (RE) edges of the t-DNA. Expression of the NPTII gene is under the control of the promoter and of the terminator of *A. tumefaciens* nopalin synthase. Downstream, the β-glucuronidase gene which has been cloned between the unique *Bam*H1 and the unique *Sac*1 sites, is under the control of the cauliflower mosaic virus (CaMV) 35S promoter and nopalin synthase gene polyadenylation signal from the Ti plasmid. The gene encoding the SAT is inserted between the *BamH1* and *Sac1* sites of the vector, from which has been deleted the β-glucuronidase gene (**Figure** 12).

Transgene expression in seeds

A construct similar to that shown in **Figures 11** or **12** is prepared with the aim of obtaining specific expression of the transgene in the seeds. This strategy may be important since seeds make up the main contribution to the animal diet. For this, the constitutive tobacco mosaic promoter is replaced with a promoter which allows specific expression of the transgene during the setting up of the seeds' stocks.

Example 9. Transformation of tobacco

Young leaves of tobacco plants (aged from 3 to 4 weeks) whose surface is sterilized with a 10% (V/V) solution of bleach for 10 min then rinsed with sterile water, are cut up with a punch (30 discs per construct). 20 ml of a 48-hour culture of *Agrobacterium tumefaciens* EHA105 (containing the pBI121 vector modified according to the invention) are centrifuged and then resuspended in 4 ml of a 10 mM solution of $MgSO_4$. The foliar discs are incubated for a few minutes in the solution of agrobacteria, then transferred to MS medium (Sigma M-5519) supplemented with 0.05 mg/l of α-naphthaleneacetic acid (NAA, Sigma), 2 mg/l of 6-benzylaminopurine (BAP) and 7 mg/l of phytoagar, for 2 to 3 days. The foliar discs are then transferred to an identical medium to which are added 350 mg/l of cefotaxin (bacteriostatic) and 75 mg/l of kanamycin (selection agent). After 2 weeks, discs on which have

developed calli as well as young shoots, are subcultured in identical medium in order to accelerate growth of the shoots. A week later, the green shoots are excised and transferred into the same medium, without hormone, in order to allow the development of roots, this for about 2 weeks, at the end of which the young plants are transferred into earth and cultivated in a hothouse.

Example 10. Analysis of results for SAT3 and SAT1' (L78443) (truncated form of the SAT1 U22964) transgenic plants and controls

The impact of the expression of SAT3, SAT1' or OTP-SAT3 in leaves or in seeds of tobacco plants is analysed as regards the content of sulphur compounds; cysteine, methionine (and derivatives such as S-methylmethionine or SMM) and glutathione. The cysteine and glutathione are evidenced by the method of Fahey ([33] Fahey, R.C. and Newton, G.L. Methods Enzymol. (1987) 143, 85-96), after derivatization of the compounds by thiolyte (mBBR from Calbiochem) and separation by high performance liquid chromatography (HPLC) [33]. The free methionine and SMM are assayed by the methods for assaying free amino acids after extraction, derivatization with ortho-phthalaldehyde, and separation by HPLC ([34] Brunet, P. et al., J. Chrom. (1988) 455, 173-182). The serine acetyltransferase activity is measured as described in

the methodology for assay of formed *O*-acetylserine, by the HPLC technique, or by the method of coupling in the presence of an excess of *O*-acetylserine (thiol) lyase [12], [14]. The SAT transgene activity in transformed plants (i.e. *in vivo*) is revealed by assaying the *O*-acetylserine, which is produced during activity of the enzyme and is transiently accumulated in the cell.

The *O*-acetylserine in the plant extracts is assayed following the protocol below.

After crushing tobacco leaves to a fine powder in liquid nitrogen, the extracts are taken up in 0.1 M hydrochloric acid (1 ml/100 mg of powder). After an incubation period of about 10 min, the debris is eliminated by centrifugation for 15 min at 15,000 g. A fraction of the obtained supernatant, containing the free amino acids, is derivatized for 1 min at 25°C in the presence of a solution of ortho-phthalaldehyde (solution containing 54 mg of ortho-phthalaldehyde, 10% methanol, 90% sodium borate, 400 mM, pH 9.5, and 0.2 ml of β-mercaptoethanol). The OPA-amino acid derivatives are separated by reverse phase chromatography on a UPHDO-15M column (0.46 × 150 mm - Interchim) connected to an HPLC system (Waters). The buffers used to carry out the elution are, buffer A: 85 mM sodium acetate, pH 4.5 supplemented with acetonitrile to 6% final; buffer B: 60% acetonitrile in water. Separation of the derivatives is carried out according to the gradient (1 ml/min): 0 min, 30% B in A; 8 min, 60% B in A;

9 min, 80% B in A; 10 min, 100% B; 12 min, 100% B. At the column exit, the fluorescence emitted by the derivatives is measured at 455 nm after excitation at 340 nm (SFM25 fluorimeter, Kontron).

The retention time of *O*-acetylserine under our experimental conditions is 9.5 min. The identity of the peak corresponding to *O*-acetylserine is confirmed by co-elution with a known quantity of the pure product. Moreover, a second control is carried out to confirm the position of *O*-acetylserine in the various analyses. The samples, before incubation with OPA, are treated with NaOH at a final concentration of 0.2 M. Under these conditions, the acetate group in the OH position on serine is transferred to the amine group, thus allowing the formation of *N*-acetylserine. This latter compound is no longer detected under our experimental conditions and thus leads to the disappearance of the peak which initially corresponded to *O*-acetylserine.

Plants transformed with an SAT transgene were preselected with kanamycin, and run to seed. Control plants (PBI, three independent lines which contain the transforming vector and a GUS cassette) are treated in an identical way. Analyses of the plants comprise: **1**; demonstration of insertion of the transgene into the genome by PCR, using the 5′ primer and the 3′ primer which correspond to the SAT which is used for the transformation; **2**, demonstration of the messenger by

analysis of messengers using probes which correspond to the SAT transgenes used for transforming the plants according to known techniques; **3**, demonstration of enzyme activity associated with SAT protein according to methods described in the literature [14], and demonstration of transgene localization; **4**, assay of the product of the SAT reaction, i.e. *O*-acetylserine (OAS), in transformed plants; **5**, assay of cysteine and its direct derivatives, of glutathione and of methionine (and its methylated derivatives); **5**, analysis of total amino acid composition of the plants and seeds which are associated with each of the transgenes obtained (free amino acids and amino acids linked to proteins), according to traditional techniques; **6**, analysis of the impact of overexpressing SAT activity in plant cells, on the amount of enzyme activity which is associated with the sequence of assimilation of sulphur (sulphate transporters, ATP-sulphurylase, APS reductase, sulphite reductase and in particular *O*-acetylserine (thiol) lyase, the enzyme which is directly associated with SAT activity in cysteine synthesis [14]. Moreover, the enzymes associated with the synthetic pathway of methionine and the synthetic pathway of glutathione, are analysed in order to understand the impact of the cysteine content on the metabolism associated with glutathione synthesis and methionine synthesis.

Expression of the *Arabidopsis thaliana* serine acetyltransferase gene in tobacco leads to an increase in the level of cysteine, the level of glutathione and the level of methionine in tissues of transformed plants, compared to control plants. In general, this increase in the amount of free sulphur compounds is associated with transgene expression in the plant cell (**Figure** 13). Measurement is carried out on leaves from 3 different plants for each homozygous line. The SAT activity is measured as its capacity to promote cysteine synthesis, according to the protocol described above [14].

Expression of the transgene under the control of the constitutive CaMV promoter, causes the SAT capacity (maximum potential enzyme activity measured *in vitro*) to increase by a factor of 2 to 8, compared to the level measured in control plants (plants transformed with an empty vector). To determine the real activity of the SAT transgene, the amount of *O*-acetylserine (free OAS) was measured. Thus, it was possible to multiply the level of OAS in plant cells (average level of 4 nmol/g of fresh material for control plants, 6 independent measurements) by a factor of 2 to 10, in transformed plants (2 independent measurements). Thus, for most SAT transgenes, associated with the clear increase in the capacity of SAT enzyme activity, is an increase in free intracellular OAS which results from the transgene

activity *in vivo*, and an increase in the amount of free cysteine, compared to control plants (**Figure** 14). The cysteine content in the control plants (PBI) and in the T2 tobacco plants transformed with an SAT (SAT1' and SAT3 lines), is determined as monobromobimane derivatives, by HPLC, for 3 plants per line [33]. The cysteine content of the transgenic lines is increased 2- to 10-fold in comparison with control plants (PBI).

The amount of free cysteine in most transgenic plants which express an SAT is significantly higher, 2 to 10-fold, than the natural level which is measured in control plants PBI (of a value of 5 nmol/g of fresh material, average calculated from three independent lines, each containing 5 plants). This impact of SAT expression is observed as early as the T1 generation. On the other hand, no correlation could be seen between amount of cysteine (and moreover of free OAS) and the SAT activity from transgenes which are measured *in vitro*. On the other hand, a significant positive correlation could be measured between amount of cellular OAS and cysteine level in the cell (**Figure** 15). *In vivo*, a 3- to 10-fold increase, compared to control plants, in the level of free *O*-acetylserine, which is linked to transgene activity, results in a 3- to 8-fold increase in the level of cysteine in the plants. Analysis was carried out on fully developed leaves (about 2 months) of plants homozygous for the transgene. The control plants are

plants transformed with empty constructs (PBI). An increase in the amount of free cellular OAS which is linked to SAT transgene activity in transformed plants, correlates positively with increase in the amount of cysteine. Thus, an average 6-fold increase in the level of free OAS is associated with a 6-fold increase in the level of cysteine. The slope associated with the distribution of the points is 1.06 +/- 0.09 (coefficient of regression 0.67). It indicates that for each molecule of OAS accumulated, one mole of cysteine is synthesized. The value of this slope and the absence of a plateau observed under our experimental conditions, indicate the sulphide formation (assimilation of sulphate and reduction to sulphide) is not a limiting pathway and that SAT activity seems to be the limiting factor in the cell for cysteine formation (**Figure** 1).

The subcellular localization of the SAT1' (truncated form of SAT1) transgene and the SAT3 transgene in transformed tobacco plants was made clear by preparation of the chloroplast fraction of transformed plants which present the highest enzyme activity, compared with PBI plants (controls). The activity associated with the chloroplast compartment is compared with that measured in the total extract (**Figure** 16). The values for serine acetyltransferase activity correspond to 3 lines for the PBI plants (5 plants per line), to 5 lines for SAT1' and SAT3, each

being represented by 5 plants. The columns in grey correspond to the activities measured in the total extract from each of the lines, and the columns in black represent the average of the activities measured in each of the chloroplast preparations.

These results establish definitively that SAT3 is an isoform of the serine acetyltransferase located in the cytosol of plant cells, and that the truncated form of SAT1 (absence of transit peptide) is also located in the cytosolic compartment. With regard to SAT3, these results confirm our interpretations which are derived from analysis of the protein sequence [12].

A direct consequence of increasing the level of cellular cysteine is increased synthesis of glutathione and methionine (see **Figure** 1). Cysteine is destined for multiple usage and besides its incorporation into proteins, and its participation in the synthesis of multiple compounds, such as vitamins (biotin, thiamine, etc. and other sulphur derivatives in the cell), cysteine also participates in the synthesis of glutathione (tripeptide which is associated with numerous plant defence mechanisms and which is considered to be a reservoir for cysteine) and of methionine. Specifically in plants which are transformed with the SAT transgene, the level of glutathione correlates directly with that of cysteine, and is reflected by an increase of 2 to 7 times the

natural level which is measured in control plants (PBI) (**Figure** 17). The correlation coefficient which is calculated for the distribution of the points is 0.92. A 4-fold increase in cysteine content in transgenic tobacco plants which overexpress SAT results in a 3- to 4-fold increase in the level of glutathione. Analysis was carried out using fully developed leaves (about 2 months) from plants homozygous for the transgene. The control plants are plants which are transformed with empty constructs.

This result indicates that cysteine is the limiting factor in glutathione synthesis in the plant cell. Thus, indirectly, the consequence of any modification of the level of serine acetyltransferase in the cell, will be to increase the amount of intracellular glutathione, by increasing the level of cysteine. This result implies that the transgenic plants obtained have acquired properties of stress resistance compared to the control plants (PBI). This aspect was observed recently ([34] Blaszczyl A. et al., 1999, The Plant Journal 20, 237-243). Moreover, the amount of cysteine and of glutathione which is considered to be a reservoir, brings about increased availability at the time of synthesis of polypeptides rich in cysteine (for example for resistance to phytopathogenic attack), and rich in cysteine and in methionine (for animal foods).

An increase in cysteine in the plant cell also brings about an increase in the relative amount of methionine (**Figure** 18). On the other hand, unlike the results observed for glutathione, the curve has a plateau, which seems to indicate the existence of another control site which would limit methionine synthesis. Moreover, homocysteine, which is derived from the trans-sulphuration pathway, and is the sulphur precursor in cysteine synthesis, does not seem to accumulate. This observation thus indicates that the folate pool in the plant cell, which is essential for methylation and for methionine formation, is not a limiting factor. This limitation would thus be situated downstream of cysteine and upstream of homocysteine. It concerns the synthesis of the carbon precursor for the aspartate-derived methionine synthesis (*O*-phosphohomoserine and/or cystathionine). The level of aspartokinase (the first enzyme of the aspartate pathway for the synthesis of lysine, threonine and methionine) is controlled by several effectors, such as threonine and *S*-adenosylmethionine (SAM) which comes from methionine synthesis [3]. Cystathionine γ-synthase (see **Figure** 1) is directly regulated at the transcriptional level [3] and, more exactly, methionine or one of its derivatives controls the stability of its messenger [4]. The maximum plateau which is obtained under our experimental conditions is of the order of 39 +/- 7 nmol of methionine/g of fresh material, which

corresponds to a multiplication of the average natural level which is of the order of 6 +/- 2 nmol per g of fresh material (PBI control). The maximum value which is obtained for methionine requires an increase in the amount of cysteine in the cell of 4 to 5 times its maximum level. The regression coefficient is 0.50.

Moreover, an increase in the methionine in the cells causes the level of *S*-methylmethionine (SMM) to multiply from 2- to 10-fold, according to the plant. SMM is derived directly from the methylation of methionine in the presence of *S*-adenosylmethionine. This compound is important to the cell, and is a form of transport of methyl groups (of methionine) in the plant. In the presence of one molecule of homocysteine (the sulphur precursor in methionine synthesis, and which is derived from cysteine), SMM allows the synthesis of two molecules of methionine ([3], [35], Bourgis et al., 1999, Plant Cell 11, 1485-1497). It may thus have a primordial role at the time of storage protein synthesis in the seed. Moreover, SMM is the direct precursor for the synthesis of compounds such as 3-dimethylsulphoniopropionate which is involved in the resistance of plants to salt stress ([36] Hanson A.D. et al., 1994, Plant Physiol. 105, 103-110). Such an approach has many consequences, in particular for increasing the potentialities of plants on grounds rich in salt.

Evidence for a regulatory role in the sulphate assimilation pathway *in vivo*.

Serine acetyltransferase is taken to be a limiting factor for the assimilation of sulphur and for the synthesis of cysteine. Its role in bacteria is important since the reaction product, (*O*-acetylserine, OAS) or its derivative (*N*-acetylserine), is the effector which modulates the expression of the genes of the sequence of assimilation of sulphur, such as: **1**, sulphate transport, **2**, ATP sulphurylase, **3**, APS kinase, and **4**, PAPS reductase ([37] Kredich N.M., 1987, in *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, pp. 419-428). In plants, a role has been shown for OAS in modulating the expression of several genes, which concerns sulphate transporters, ([38] Smith F.W. et al., 1997, The Plant Journal 12, 875-884; [39] Hawkesford M.J. et al. 1995, Z. Pfanzenernärh. Bodenk. 158, 55-57; [40] Clarkson D.T. et al. 1999, Plant Physiol. Biochem. 37, 283-290), ATP sulphurylase [39-40] and APS reductase ([41] Neuenschwander U. et al. 1991, Plant Physiol., 97, 253-258). The role of serine acetyltransferase activity in gene modulation has been proposed based on the kinetics of the cysteine synthase complex (bienzyme complex composed of serine acetyltransferase and of *O*-acetylserine (thiol) lyase) ([41] Droux et al. in Sulphur and Nutrition in Plants, in press), and has led to the description of a model to describe the mechanism

of gene regulation. The role of OAS is also determinant in the regulation of gene expression during seed formation ([42] Kim H. et al., 1999, Planta 209, 282-289).

In transgenic plants which overexpress an SAT in the cytosol, a transient increase in OAS was shown (increase of 2 to 10 times its natural level, see **Figure** 15). In parallel, in most transgenic plants, an increase in OASTL activity was measured (**Figure** 19). This increase of 2 to 5 times compared to the activity which is measured in PBI controls, concerns only the chloroplast-associated activity. Moreover, in a Western Blot, the signal which is observed is stronger in most transgenic lines (**Figure** 20), indicating that the increase in activity corresponds to an induction of *de novo* synthesis of OASTL. This original result corresponds to the first demonstration of the role of OAS (*in planta*) in the modulation of genes of the sulphate assimilation pathway, in particular for chloroplast OASTL.

Referring to **Figure** 20, an equivalent amount of protein (0.150 mg) undergoes SDS-PAGE (12%), and after separation, the proteins are transferred onto a nitrocellulose membrane. The presence of OASTL is revealed by incubation with antibodies which have been raised against chloroplast OASTL from spinach leaves [7].

Overexpression of SAT in plant cells thus causes the capacity to synthesize cysteine in the chloroplast to increase. It can, therefore, be assumed that the expression of genes encoding enzymes of the sulphate assimilation and reduction pathway (sulphate transporter, ATP sulphurylase, APS reductase, sulphite reductase) is also modulated like OASTL (and references [38-41]).

The increase in the intracellular content of OAS (which is derived from SAT activity) signals a state of artificial sulphur stress (absence of sufficient reduced sulphur) in the cell (in transformed plants), which leads to induction of the enzymes of the sulphate assimilation pathway.

Impact of increasing cysteine in a cell on the general content of amino acids. This increase in sulphur compounds is accompanied by an increase in the content of essential amino acids, such as threonine, isoleucine and lysine (their amount is doubled, on average). On the other hand, the level of glutamate is halved, as is that of aspartate. This latter observation is directly linked to the increase in the amount of THR, LYS and ILE. All the increases in amino acids correlate with an increase in serine acetyltransferase (SAT3 or SAT1') activity in the cytosol. Moreover, an increase in these sulphur compounds leads to an improvement in the nutritional ratio N/S of the plant (on the basis of

free amino acids). It is reflected by a drop in this relative ratio, due to the enrichment in total sulphur compounds (cysteine, methionine, SMM and glutathione). This factor is important since it conditions the polypeptide content of the seeds, and leads to enrichment (or impoverishment if the N/S ratio is too high) of storage proteins whch are rich in sulphur-containing amino acids, to the detriment of polypeptides which are lacking in these compounds.

Example 11. Analysis of OTP-SAT3 (OTP-SAT1') transgenic plants

Analysis of transformants at the TO stage of transgenic plants which express a cysteine-insensitive SAT (here for example, SAT3 or SAT1'; truncated form of SAT1 U22964), in leaves or in seeds (under the control of a seed-specific promoter), reveals an increase in free cysteine content, but also in glutathione content (2.6 times the natural level), and in methionine content. Plants which express these same isoforms in the cytosol under the control of a seed-specific promoter show a level of sulphur compounds which is higher that in control plants.

Example 12. Analysis of results for SAT1 (cDNA U22964 or SAT1jw, transit peptide form) transgenic plants and control plants.

The impact of expression of serine acetyltransferase in mitochondria was analysed by transforming plants with the construct (**Figure** 12) which contains the entire SAT1 sequence. Analysis of plants at the TO stage makes it possible to show an increase in free cysteine in the cell (**Figure** 21). Analysis is carried out on leaves which are formed before appearance of the floral scape. The fourteen lines show a 2- to 6-fold multiplication in cysteine level, compared with the control plant (PBI).

The increase in cysteine is accompanied by a general effect on the amount of sulphur compounds, with a 4-fold multiplication in the amount of glutathionine in the cell (**Figure** 22). Unlike the case of SAT expression in the cytosolic compartment, the general appearance of the distribution of values in the different lines, shows a plateau which would indicate limitation in glutathione synthesis. This limitation may concern the level of glutamate and/or glycine or may concern glutathione control of its own synthesis (retroinhibition of one of the enzymes which participate in glutathione synthesis, enzyme E6 and/or enzyme E7 see **Figure** 1).

Similarly, the amount of methionine is multiplied 2- to 3-fold compared to the natural level which is measured in control plants.

SEQUENCE LISTING

<110> RHONE-POULENC AGRO

<120> Method for increasing the content of cysteine, methionine and glutathione in plants, and plants obtained

<130>

<140>
<141>

<150> FR9816163
<151> 1998-12-17

<160> 17

<170> PatentIn Ver. 2.1

<210> 1
<211> 984
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS

<222>(31)..(972)

```
<400> 1
gagagaggat ctctcttcca atcataaacc atg gca aca tgc ata gac aca tgc   54
                                 Met Ala Thr Cys Ile Asp Thr Cys
                                   1               5

cga acc ggt aat acc caa gac gat gat tcc cgg ttc tgt tgc atc aag   102
Arg Thr Gly Asn Thr Gln Asp Asp Asp Ser Arg Phe Cys Cys Ile Lys
     10                  15                  20

aat ttc ttt cga ccc ggt ttc tct gta aac cgg aag att cac cac acc   150
Asn Phe Phe Arg Pro Gly Phe Ser Val Asn Arg Lys Ile His His Thr
 25                  30                  35                  40

caa atc gaa gat gac gat gat gtc tgg atc aag atg ctt gaa gaa gcc   198
Gln Ile Glu Asp Asp Asp Asp Val Trp Ile Lys Met Leu Glu Glu Ala
                 45                  50                  55

aaa tcc gat gtt aaa caa gaa ccc att tta tca aac tac tac tac gct   246
Lys Ser Asp Val Lys Gln Glu Pro Ile Leu Ser Asn Tyr Tyr Tyr Ala
             60                  65                  70

tcg atc aca tct cat cga tct tta gag tct gct tta gct cac atc ctc   294
Ser Ile Thr Ser His Arg Ser Leu Glu Ser Ala Leu Ala His Ile Leu
         75                  80                  85

tcc gta aag ctc agc aat tta aac cta cca agc aac aca ctc ttc gaa   342
Ser Val Lys Leu Ser Asn Leu Asn Leu Pro Ser Asn Thr Leu Phe Glu
     90                  95                 100

ctg ttc ata agc gtt tta gaa gaa agc cct gag atc atc gaa tcc acg   390
Leu Phe Ile Ser Val Leu Glu Glu Ser Pro Glu Ile Ile Glu Ser Thr
105                 110                 115                 120

aag caa gat ctt ata gca gtc aaa gaa aga gac cca gct tgt ata agc   438
Lys Gln Asp Leu Ile Ala Val Lys Glu Arg Asp Pro Ala Cys Ile Ser
                125                 130                 135
```

```
tac gtt cat tgc ttc ttg ggc ttc aaa ggc ttc ctc gct tgt caa gct     486
Tyr Val His Cys Phe Leu Gly Phe Lys Gly Phe Leu Ala Cys Gln Ala
            140                 145                 150

cat cga ata gct cat acc ctc tgg aaa cag aac cga aaa atc gta gct     534
His Arg Ile Ala His Thr Leu Trp Lys Gln Asn Arg Lys Ile Val Ala
        155                 160                 165

tta ttg atc caa aac aga gta tca gaa tct ttc gcc gtc gat att cat     582
Leu Leu Ile Gln Asn Arg Val Ser Glu Ser Phe Ala Val Asp Ile His
    170                 175                 180

ccc gga gcg aag atc gga aaa ggg att ctt tta gac cat gcg acg ggc     630
Pro Gly Ala Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly
185                 190                 195                 200

gtg gtg atc gga gag acg gcg gtg gtt gga gac aat gtc tcg att cta     678
Val Val Ile Gly Glu Thr Ala Val Val Gly Asp Asn Val Ser Ile Leu
                205                 210                 215

cac gga gtg acc ttg gga gga aca ggg aaa cag agt ggt gat cgg cat     726
His Gly Val Thr Leu Gly Gly Thr Gly Lys Gln Ser Gly Asp Arg His
            220                 225                 230

ccg aag att ggt gat ggt gtg ttg att gga gct ggg agt tgt ata ttg     774
Pro Lys Ile Gly Asp Gly Val Leu Ile Gly Ala Gly Ser Cys Ile Leu
        235                 240                 245

ggg aat ata aca atc ggt gag gga gct aag att gga tca ggg tcg gtg     822
Gly Asn Ile Thr Ile Gly Glu Gly Ala Lys Ile Gly Ser Gly Ser Val
    250                 255                 260

gtg gtt aag gat gtg ccg gcg cgt acg acg gcg gtt gga aat ccg gcg     870
Val Val Lys Asp Val Pro Ala Arg Thr Thr Ala Val Gly Asn Pro Ala
265                 270                 275                 280

agg ttg att ggt ggg aaa gag aat ccg aga aag cat gat aag att cct     918
Arg Leu Ile Gly Gly Lys Glu Asn Pro Arg Lys His Asp Lys Ile Pro
                285                 290                 295

tgt ctg act atg gac cag aca tcg tat tta acc gag tgg tct gat tat     966
Cys Leu Thr Met Asp Gln Thr Ser Tyr Leu Thr Glu Trp Ser Asp Tyr
            300                 305                 310

gtg att taacacaaat gt                                               984
Val Ile
```

<210> 2
<211> 974
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (31)..(966)

<400> 2

```
gagagaggat cctcttatcg ccgcgttaat atg cca ccg gcc gga gaa ctc cga      54
                                 Met Pro Pro Ala Gly Glu Leu Arg
                                 1               5

cat caa tct cca tca aag gag aaa cta tct tcc gtt acc caa tcc gat     102
His Gln Ser Pro Ser Lys Glu Lys Leu Ser Ser Val Thr Gln Ser Asp
    10                  15                  20
```

gaa gca gaa gca gcg tca gca gcg ata tcc gcg gca gct gca gat gcg 150
Glu Ala Glu Ala Ala Ser Ala Ala Ile Ser Ala Ala Ala Ala Asp Ala
25 30 35 40

gaa gct gcc gga tta tgg aca cag atc aag gcg gaa gct cgc cgt gat 198
Glu Ala Ala Gly Leu Trp Thr Gln Ile Lys Ala Glu Ala Arg Arg Asp
45 50 55

gct gag gcg gag cca gct tta gct agc tat cta tat tcg acg att ctt 246
Ala Glu Ala Glu Pro Ala Leu Ala Ser Tyr Leu Tyr Ser Thr Ile Leu
60 65 70

tct cat tcg tct ctt gaa cga tct atc tcg ttc cat cta gga aac aag 294
Ser His Ser Ser Leu Glu Arg Ser Ile Ser Phe His Leu Gly Asn Lys
75 80 85

ctt tgt tcc tca acg ctt tta tcc aca ctt tta tac gat ctg ttc tta 342
Leu Cys Ser Ser Thr Leu Leu Ser Thr Leu Leu Tyr Asp Leu Phe Leu
90 95 100

aac act ttt tcc tcc gat cct tct ctt cgt aac gcc acc gtt gca gat 390
Asn Thr Phe Ser Ser Asp Pro Ser Leu Arg Asn Ala Thr Val Ala Asp
105 110 115 120

cta cgt gct gct cgt gtt cgt gat cct gct tgt atc tcg ttc tct cat 438
Leu Arg Ala Ala Arg Val Arg Asp Pro Ala Cys Ile Ser Phe Ser His
125 130 135

tgt ctc ctc aat tac aaa ggc ttc tta gct att cag gcg cat cgt gta 486
Cys Leu Leu Asn Tyr Lys Gly Phe Leu Ala Ile Gln Ala His Arg Val
140 145 150

tca cac aag cta tgg aca caa tca cgg aag cca tta gca tta gct cta 534
Ser His Lys Leu Trp Thr Gln Ser Arg Lys Pro Leu Ala Leu Ala Leu
155 160 165

cac tca aga atc tcc gat gta ttc gct gtt gat atc cat cca gca gcg 582
His Ser Arg Ile Ser Asp Val Phe Ala Val Asp Ile His Pro Ala Ala
170 175 180

aag atc gga aaa ggg ata ctt cta gac cac gca acc gga gtt gta gtc 630
Lys Ile Gly Lys Gly Ile Leu Leu Asp His Ala Thr Gly Val Val Val
185 190 195 200

gga gaa aca gcg gtg att ggg aac aat gtt tca atc ctt cac cat gtg 678
Gly Glu Thr Ala Val Ile Gly Asn Asn Val Ser Ile Leu His His Val
205 210 215

aca cta ggt gga aca ggt aaa gct tgt gga gat aga cat ccg aag atc 726
Thr Leu Gly Gly Thr Gly Lys Ala Cys Gly Asp Arg His Pro Lys Ile
220 225 230

ggt gac ggt tgt ttg att gga gct gga gcg act att ctt gga aat gtg 774
Gly Asp Gly Cys Leu Ile Gly Ala Gly Ala Thr Ile Leu Gly Asn Val
235 240 245

aag att ggt gca ggt gct aaa gta gga gct ggt tct gtt gtg ctg att 822
Lys Ile Gly Ala Gly Ala Lys Val Gly Ala Gly Ser Val Val Leu Ile
250 255 260

gac gtg cct tgt cga ggt act gcg gtt ggg aat ccg gcg aga ctt gtc 870
Asp Val Pro Cys Arg Gly Thr Ala Val Gly Asn Pro Ala Arg Leu Val
265 270 275 280

gga ggg aaa gag aag cca acg att cat gat gag gaa tgt cct gga gaa 918
Gly Gly Lys Glu Lys Pro Thr Ile His Asp Glu Glu Cys Pro Gly Glu
285 290 295

tcg atg gat cat act tca ttc atc tcg gaa tgg tca gat tac atc ata 966
Ser Met Asp His Thr Ser Phe Ile Ser Glu Trp Ser Asp Tyr Ile Ile
300 305 310

taaagttg 974

<210> 3
<211> 1048
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (31)..(1038)

<400> 3
gagagaggat cccctcctcc tcctcctcct atg gct gcg tgc atc gac acc tgc 54
Met Ala Ala Cys Ile Asp Thr Cys
1 5

cgc act ggt aaa ccc cag att tct cct cgc gat tct tct aaa cac cac 102
Arg Thr Gly Lys Pro Gln Ile Ser Pro Arg Asp Ser Ser Lys His His
10 15 20

gac gat gaa tct ggc ttt cgt tac atg aac tac ttc cgt tat cct gat 150
Asp Asp Glu Ser Gly Phe Arg Tyr Met Asn Tyr Phe Arg Tyr Pro Asp
25 30 35 40

cga tct tcc ttc aat gga acc cag acc aaa acc ctc cat act cgt cct 198
Arg Ser Ser Phe Asn Gly Thr Gln Thr Lys Thr Leu His Thr Arg Pro
45 50 55

ttg ctt gaa gat ctc gat cgc gat gct gaa gtt gat gat gtt tgg gcc 246
Leu Leu Glu Asp Leu Asp Arg Asp Ala Glu Val Asp Asp Val Trp Ala
60 65 70

aaa atc cga gaa gag gct aaa tct gat atc gcc aaa gaa cct att gtt 294
Lys Ile Arg Glu Glu Ala Lys Ser Asp Ile Ala Lys Glu Pro Ile Val
75 80 85

tcc gct tat tat cac gct tcg att gtt tct cag cgt tcg ttg gaa gct 342
Ser Ala Tyr Tyr His Ala Ser Ile Val Ser Gln Arg Ser Leu Glu Ala
90 95 100

gcg ttg gcg aat act tta tct gtt aaa ctc agc aat ttg aat ctt cca 390
Ala Leu Ala Asn Thr Leu Ser Val Lys Leu Ser Asn Leu Asn Leu Pro
105 110 115 120

agc aac acg ctt ttc gat ttg ttc tct ggt gtt ctt caa gga aac cca 438
Ser Asn Thr Leu Phe Asp Leu Phe Ser Gly Val Leu Gln Gly Asn Pro
125 130 135

gat att gtt gaa tct gtc aag cta gat ctt tta gct gtt aag gag aga 486
Asp Ile Val Glu Ser Val Lys Leu Asp Leu Leu Ala Val Lys Glu Arg
140 145 150

gat cct gct tgt ata agc tac gtt cat tgt ttc ctt cac ttt aaa ggc 534
Asp Pro Ala Cys Ile Ser Tyr Val His Cys Phe Leu His Phe Lys Gly
155 160 165

ttc ctc gct tgt caa gcg cat cgt att gct cat gag ctt tgg act cag 582
Phe Leu Ala Cys Gln Ala His Arg Ile Ala His Glu Leu Trp Thr Gln
170 175 180

gac aga aaa atc cta gct ttg ttg atc cag aac aga gtc tct gaa gcc 630
Asp Arg Lys Ile Leu Ala Leu Leu Ile Gln Asn Arg Val Ser Glu Ala
185 190 195 200

ttc gct gtt gat ttc cac cct gga gct aaa atc ggt acc ggg att ttg 678
Phe Ala Val Asp Phe His Pro Gly Ala Lys Ile Gly Thr Gly Ile Leu
205 210 215

cta gac cat gct acg gct att gtg atc ggt gag acg gcg gtt gtg ggg 726
Leu Asp His Ala Thr Ala Ile Val Ile Gly Glu Thr Ala Val Val Gly
220 225 230

aac aat gtt tcg att ctc cat aac gtt acg ctt gga gga acg ggg aaa 774
Asn Asn Val Ser Ile Leu His Asn Val Thr Leu Gly Gly Thr Gly Lys
235 240 245

cag tgt gga gat agg cac ccg aag att ggc gat ggg gtt ttg att gga 822
Gln Cys Gly Asp Arg His Pro Lys Ile Gly Asp Gly Val Leu Ile Gly
250 255 260

gct ggg act tgt att ttg ggg aat atc acg att ggt gaa gga gct aag 870
Ala Gly Thr Cys Ile Leu Gly Asn Ile Thr Ile Gly Glu Gly Ala Lys
265 270 275 280

att ggt gcg ggg tcg gtg gtg ttg aaa gac gtg ccg ccg cgt acg acc 918
Ile Gly Ala Gly Ser Val Val Leu Lys Asp Val Pro Pro Arg Thr Thr
285 290 295

gct gtt gga aat ccg gcg agg ttg ctt ggt ggt aaa gat aat ccg aaa 966
Ala Val Gly Asn Pro Ala Arg Leu Leu Gly Gly Lys Asp Asn Pro Lys
300 305 310

acg cat gac aag att cct ggt ttg act atg gac cag acg tcg cat ata 1014
Thr His Asp Lys Ile Pro Gly Leu Thr Met Asp Gln Thr Ser His Ile
315 320 325

tcc gag tgg tcg gat tat gta att tgaaaaagtc 1046
Ser Glu Trp Ser Asp Tyr Val Ile
330 335

<210> 4
<211> 1213
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (31)..(1203)

<220>
<221> sig_peptide
<222> (31)..(219)

<400> 4
gagagaggat ccggccgaga aaaaaaaaaa atg ttg ccg gtc aca agt cgc cgc 54
Met Leu Pro Val Thr Ser Arg Arg
1 5

cac ttc aca atg tcc cta tat atg ctc cgt tca tct tct cca cac atc 102
His Phe Thr Met Ser Leu Tyr Met Leu Arg Ser Ser Ser Pro His Ile
10 15 20

```
aat cat cac tct ttc ctt ctt cct tct ttt gtt tcc tcc aaa ttc aaa   150
Asn His His Ser Phe Leu Leu Pro Ser Phe Val Ser Ser Lys Phe Lys
 25                 30                  35                  40

cac cat act tta tct cct cct cct tct cct cct cct cct cct cct atg   198
His His Thr Leu Ser Pro Pro Pro Ser Pro Pro Pro Pro Pro Pro Met
                 45                  50                  55

gct gcg tgc atc gac acc tgc cgc act ggt aaa ccc cag att tct cct   246
Ala Ala Cys Ile Asp Thr Cys Arg Thr Gly Lys Pro Gln Ile Ser Pro
             60                  65                  70

cgc gat tct tct aaa cac cac gac gat gaa tct ggc ttc cgt tac atg   294
Arg Asp Ser Ser Lys His His Asp Asp Glu Ser Gly Phe Arg Tyr Met
         75                  80                  85

aac tac ttc cgt tat cct gat cga tct tcc ttc aat gga acc cag acc   342
Asn Tyr Phe Arg Tyr Pro Asp Arg Ser Ser Phe Asn Gly Thr Gln Thr
     90                  95                 100

aaa acc ctc cat act cgt cct ttg ctt gaa gat ctc gat cgc gac gct   390
Lys Thr Leu His Thr Arg Pro Leu Leu Glu Asp Leu Asp Arg Asp Ala
105                 110                 115                 120

gaa gtc gat gat gtt tgg gcc aaa atc cga gaa gag gct aaa tct gat   438
Glu Val Asp Asp Val Trp Ala Lys Ile Arg Glu Glu Ala Lys Ser Asp
                125                 130                 135

atc gcc aaa gaa cct att gtt tcc gct tat tat cac gct tcg atc gtt   486
Ile Ala Lys Glu Pro Ile Val Ser Ala Tyr Tyr His Ala Ser Ile Val
            140                 145                 150

tct cag cgt tcg ttg gaa gct gcg ttg gcg aat act tta tct gtt aaa   534
Ser Gln Arg Ser Leu Glu Ala Ala Leu Ala Asn Thr Leu Ser Val Lys
        155                 160                 165

ctc agc aat ttg aat ctt cca agc aac acg ctt ttc gat ttg ttc tct   582
Leu Ser Asn Leu Asn Leu Pro Ser Asn Thr Leu Phe Asp Leu Phe Ser
    170                 175                 180

ggt gtt ctt caa gga aac cca gat att gtt gaa tct gtc aag cta gat   630
Gly Val Leu Gln Gly Asn Pro Asp Ile Val Glu Ser Val Lys Leu Asp
185                 190                 195                 200

ctt tta gct gtt aag gag aga gat cct gct tgt ata agc tac gtt cat   678
Leu Leu Ala Val Lys Glu Arg Asp Pro Ala Cys Ile Ser Tyr Val His
                205                 210                 215

tgt ttc ctt cac ttt aaa ggc ttc ctc gct tgt caa gcg cat cgt att   726
Cys Phe Leu His Phe Lys Gly Phe Leu Ala Cys Gln Ala His Arg Ile
            220                 225                 230

gct cat gag ctt tgg acc cag gac aga aaa atc cta gct ttg ttg atc   774
Ala His Glu Leu Trp Thr Gln Asp Arg Lys Ile Leu Ala Leu Leu Ile
        235                 240                 245

cag aac aga gtc tct gaa gcc ttc gct gtt gat ttc cac cct gga gct   822
Gln Asn Arg Val Ser Glu Ala Phe Ala Val Asp Phe His Pro Gly Ala
    250                 255                 260

aaa atc ggt acc ggg att ttg cta gac cat gct acg gct att gtg atc   870
Lys Ile Gly Thr Gly Ile Leu Leu Asp His Ala Thr Ala Ile Val Ile
265                 270                 275                 280
```

ggt gag acg gcg gtt gtg ggg aac aat gtt tcg att ctc cat aac gtt 918
Gly Glu Thr Ala Val Val Gly Asn Asn Val Ser Ile Leu His Asn Val
285 290 295

acg ctt gga gga acg ggg aaa cag tgt gga gat agg cac ccg aag att 966
Thr Leu Gly Gly Thr Gly Lys Gln Cys Gly Asp Arg His Pro Lys Ile
300 305 310

ggc gat ggg gtt ttg att gga gct ggg act tgt att ttg ggg aat atc 1014
Gly Asp Gly Val Leu Ile Gly Ala Gly Thr Cys Ile Leu Gly Asn Ile
315 320 325

acg att ggt gaa gga gct aag att ggt gcg ggg tcg gtg gtg ttg aaa 1062
Thr Ile Gly Glu Gly Ala Lys Ile Gly Ala Gly Ser Val Val Leu Lys
330 335 340

gac gtg ccg ccg cgt acg acg gct gtt gga aat ccg gcg agg ttg ctg 1110
Asp Val Pro Pro Arg Thr Thr Ala Val Gly Asn Pro Ala Arg Leu Leu
345 350 355 360

ggt ggt aaa gat aat ccg aaa acg cat gac aag att cct ggt ttg act 1158
Gly Gly Lys Asp Asn Pro Lys Thr His Asp Lys Ile Pro Gly Leu Thr
365 370 375

atg gac cag acg tcg cat ata tcc gag tgg tcg gat tat gta att 1203
Met Asp Gln Thr Ser His Ile Ser Glu Trp Ser Asp Tyr Val Ile
380 385 390

tgaaaaagtc 1213

<210> 5
<211> 1080
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (1)..(1080)

<220>
<221> transit_peptide
<222> (1)..(96)

<400> 5
atg gtg gat cta tct tct ttt agc ctc ctt ttt gct ttc tcc gtc tct 48
Met Val Asp Leu Ser Ser Phe Ser Leu Leu Phe Ala Phe Ser Val Ser
1 5 10 15

ctc tct ttt gtc caa tca aaa aga gtt tgt gat tct tct tta tcg tct 96
Leu Ser Phe Val Gln Ser Lys Arg Val Cys Asp Ser Ser Leu Ser Ser
20 25 30

cct tgg aga gat atg aat ggc gat gag ctt cct ttc gag agt ggt ttc 144
Pro Trp Arg Asp Met Asn Gly Asp Glu Leu Pro Phe Glu Ser Gly Phe
35 40 45

gag gtt tac gct aag gga act cat aag tca gag ttt gac tcg aat ttg 192
Glu Val Tyr Ala Lys Gly Thr His Lys Ser Glu Phe Asp Ser Asn Leu
50 55 60

ctt gat cct cgt tct gat cct att tgg gat gct ata aga gaa gaa gct 240
Leu Asp Pro Arg Ser Asp Pro Ile Trp Asp Ala Ile Arg Glu Glu Ala
65 70 75 80

```
aaa ctt gag gca gag aaa gag cct att ttg agt agc ttc ttg tat gct     288
Lys Leu Glu Ala Glu Lys Glu Pro Ile Leu Ser Ser Phe Leu Tyr Ala
                85                  90                  95

ggt atc tta gca cat gat tgt tta gag caa gct tta ggg ttt gtt cta     336
Gly Ile Leu Ala His Asp Cys Leu Glu Gln Ala Leu Gly Phe Val Leu
            100                 105                 110

gcc aac cgt ctc caa aac cca acc ttg ttg gca aca caa ctc ttg gat     384
Ala Asn Arg Leu Gln Asn Pro Thr Leu Leu Ala Thr Gln Leu Leu Asp
        115                 120                 125

ata ttt tat ggt gtt atg atg cat gac aaa ggt att cag agt tcg att     432
Ile Phe Tyr Gly Val Met Met His Asp Lys Gly Ile Gln Ser Ser Ile
    130                 135                 140

cgc cat gat ctc cag gca ttt aaa gat cgt gat cct gct tgt ctg tcg     480
Arg His Asp Leu Gln Ala Phe Lys Asp Arg Asp Pro Ala Cys Leu Ser
145                 150                 155                 160

tat agt tct gct att tta cat ctg aag ggt tat cat gcg tta caa gca     528
Tyr Ser Ser Ala Ile Leu His Leu Lys Gly Tyr His Ala Leu Gln Ala
                165                 170                 175

tat agg gtt gcg cat aaa ctg tgg aat gaa ggg agg aaa cta tta gct     576
Tyr Arg Val Ala His Lys Leu Trp Asn Glu Gly Arg Lys Leu Leu Ala
            180                 185                 190

ctt gca ttg caa agc cga ata agc gag gtt ttt ggc att gac ata cat     624
Leu Ala Leu Gln Ser Arg Ile Ser Glu Val Phe Gly Ile Asp Ile His
        195                 200                 205

cca gcg gca aga att ggg gag gga ata ttg ttg gat cat gga act gga     672
Pro Ala Ala Arg Ile Gly Glu Gly Ile Leu Leu Asp His Gly Thr Gly
    210                 215                 220

gtg gtc att ggt gag acc gct gtg ata ggc aac ggt gtc tcg atc tta     720
Val Val Ile Gly Glu Thr Ala Val Ile Gly Asn Gly Val Ser Ile Leu
225                 230                 235                 240

cat ggt gtg act tta gga gga acc gga aag gaa act ggc gat cgc cac     768
His Gly Val Thr Leu Gly Gly Thr Gly Lys Glu Thr Gly Asp Arg His
                245                 250                 255

cca aag ata ggt gaa ggt gca ttg ctt gga gct tgt gtg act ata ctt     816
Pro Lys Ile Gly Glu Gly Ala Leu Leu Gly Ala Cys Val Thr Ile Leu
            260                 265                 270

ggt aac ata agc ata ggt gct gga gca atg gta gct gca ggt tca ctt     864
Gly Asn Ile Ser Ile Gly Ala Gly Ala Met Val Ala Ala Gly Ser Leu
        275                 280                 285

gtg tta aaa gac gtt cct tcg cat agt gtg gtg gct gga aat cct gca     912
Val Leu Lys Asp Val Pro Ser His Ser Val Val Ala Gly Asn Pro Ala
    290                 295                 300

aaa ctg atc agg gtc atg gaa gag caa gac ccg tct cta gca atg aaa     960
Lys Leu Ile Arg Val Met Glu Glu Gln Asp Pro Ser Leu Ala Met Lys
305                 310                 315                 320

cac gat gct act aaa gag ttc ttt cga cat gta gct gat ggt tac aaa    1008
His Asp Ala Thr Lys Glu Phe Phe Arg His Val Ala Asp Gly Tyr Lys
                325                 330                 335
```

```
ggg gca caa tct aac gga cca tca ctt tca gca gga gat aca gag aaa     1056
Gly Ala Gln Ser Asn Gly Pro Ser Leu Ser Ala Gly Asp Thr Glu Lys
            340                 345                 350

gga cac act aac agc aca tca tga                                     1080
Gly His Thr Asn Ser Thr Ser
        355                 360
```

<210> 6
<211> 900
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (1)..(900)

<220>
<221> transit_peptide
<222> (1)..(90)

<400> 11

```
atg gct tgt ata aac ggc gag aat cgt gat ttt tct tcc tcg tca tct     48
Met Ala Cys Ile Asn Gly Glu Asn Arg Asp Phe Ser Ser Ser Ser Ser
1               5                   10                  15

ttg tct tct ctt cca atg att gtc tcc cgg aac ttt tct gcc aga gac     96
Leu Ser Ser Leu Pro Met Ile Val Ser Arg Asn Phe Ser Ala Arg Asp
            20                  25                  30

gat gga gag acc ggt gac gag ttt cct ttc gag agg att ttc ccg gtt     144
Asp Gly Glu Thr Gly Asp Glu Phe Pro Phe Glu Arg Ile Phe Pro Val
        35                  40                  45

tac gct aga gga acc ctt aat ccc gtg gcc gac ccg gtt ttg ctg gat     192
Tyr Ala Arg Gly Thr Leu Asn Pro Val Ala Asp Pro Val Leu Leu Asp
    50                  55                  60

ttt acc aat tct agt tat gac cca att tgg gat tct ata aga gaa gaa     240
Phe Thr Asn Ser Ser Tyr Asp Pro Ile Trp Asp Ser Ile Arg Glu Glu
65                  70                  75                  80

gct aag ctt gag gca gaa gag gag ccg gtt ttg agt agc ttc ttg tat     288
Ala Lys Leu Glu Ala Glu Glu Glu Pro Val Leu Ser Ser Phe Leu Tyr
                85                  90                  95

gct agt atc ttg tcg cat gac tgt tta gag caa gca ttg agt ttt gtt     336
Ala Ser Ile Leu Ser His Asp Cys Leu Glu Gln Ala Leu Ser Phe Val
            100                 105                 110

cta gct aac cgt ctc caa aac cct acc ttg ttg gca act cag ctt atg     384
Leu Ala Asn Arg Leu Gln Asn Pro Thr Leu Leu Ala Thr Gln Leu Met
        115                 120                 125

gat ata ttt tgc aac gtt atg gta cat gac aga ggt att caa agc tcg     432
Asp Ile Phe Cys Asn Val Met Val His Asp Arg Gly Ile Gln Ser Ser
    130                 135                 140

att cgt ctt gat gtt cag gca ttc aaa gac aga gat cct gct tgt cta     480
Ile Arg Leu Asp Val Gln Ala Phe Lys Asp Arg Asp Pro Ala Cys Leu
145                 150                 155                 160

tcg tat agt tcg gct att tta cat ctg aag ggc tat ctt gca ctg cag     528
Ser Tyr Ser Ser Ala Ile Leu His Leu Lys Gly Tyr Leu Ala Leu Gln
                165                 170                 175
```

```
gcg tat aga gta gca cat aag ttg tgg aag caa gga aga aaa cta tta      576
Ala Tyr Arg Val Ala His Lys Leu Trp Lys Gln Gly Arg Lys Leu Leu
        180                     185                     190

gca ttg gca ctg caa agc cga gta agc gag gta aga act gct gtg ata      624
Ala Leu Ala Leu Gln Ser Arg Val Ser Glu Val Arg Thr Ala Val Ile
        195                     200                     205

ggc gac cgt gtc tca att ttg cat ggt gtg aca tta gga gga act ggg      672
Gly Asp Arg Val Ser Ile Leu His Gly Val Thr Leu Gly Gly Thr Gly
    210                     215                     220

aaa gaa acc ggt gac cgc cat cca aat ata ggc gac ggt gct ctt ctt      720
Lys Glu Thr Gly Asp Arg His Pro Asn Ile Gly Asp Gly Ala Leu Leu
225                     230                     235                     240

gga gca tgt gtg act ata ctt ggt aac att aag ata ggc gct gga gca      768
Gly Ala Cys Val Thr Ile Leu Gly Asn Ile Lys Ile Gly Ala Gly Ala
                245                     250                     255

atg gta gct gct ggt tcg ctt gtg tta aag gat gtt cct tcg cat agc      816
Met Val Ala Ala Gly Ser Leu Val Leu Lys Asp Val Pro Ser His Ser
            260                     265                     270

atg gtg gct gga aat cca gca aaa ctc atc ggg ttt gtt gat gag caa      864
Met Val Ala Gly Asn Pro Ala Lys Leu Ile Gly Phe Val Asp Glu Gln
        275                     280                     285

gat cca tct atg aca atg gag cat ggt cag tct tga                      900
Asp Pro Ser Met Thr Met Glu His Gly Gln Ser
    290                     295                     300
```

<210> 7
<211> 54
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 7
gagagaggat cctctttcca atcataaacc atggcaacat gcatagacac atgc 54

<210> 8
<211> 46
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 8
ggctcaccag actaatacac taaattgtgt ttacctcgag agagag 46

<210> 9
<211> 52
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 9
gagagaggat cctcttatcg ccgcgttaat atgccaccgg ccggagaact cc 52

<210> 10
<211> 45
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 10
gagccttacc agtctaatgt agtatatttc aacctcgaga gagag 45

<210> 11
<211> 53
<212> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 11
gagagaggat cccctcctcc tcctcctcct atggctgcgt gcatcgacac ctg 53

<210> 12
<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 12
gctcaccagc ctaatacatt aaactttttc agctcgagag agag 44

<210> 13
<211> 53
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 13
gagagaggat ccggccgaga aaaaaaaaaa atgttgccgg tcacaagtcg ccg 53

<210> 14
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:

synthetic oligonucleotide

<400> 14
gagagaggat ccgacaagtt ggcataattt atggtggatc tatcttcct 49

<210> 15
<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 15
cctgtgtgat tgtcgtgtag tactctagaa actcgagaga gag 43

<210> 16
<211> 67
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 16
gagagaggat ccgacaagtt ggcataattt atggcttgta taaacggcga gaatcgtgat 60

ttttctt 67

<210> 17
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Artificial sequence description:
synthetic oligonucleotide

<400> 17
tacctcgtac cactcagaac tctagaaact cgagagagag 40

Claims

1. Method for increasing the production of cysteine, glutathione and methionine, and of sulphur derivatives thereof, by plant cells and plants, the said method consisting in overexpressing an SAT in plant cells and plants containing the said plant cells.

2. Method according to claim 1, characterized in that the SAT which is overexpressed in plant cells is a cysteine-sensitive SAT.

Method according to claim 2, characterized in that the SAT is a plant SAT or a native SAT of bacterial origin.

4. Method according to claim 1, characterized in that the SAT which is overexpressed in plant cells is a cysteine-insensitive SAT.

5. Method according to claim 4, characterized in that the SAT is a plant SAT or an SAT of bacterial origin, or a mutated plant SAT, rendered cysteine-insensitive by mutagenesis.

6. Method according to one of claims 1 to 5, characterized in that the SAT is overexpressed in the cytoplasm of plant cells.

7. Method according to claim 6, characterized in that the SAT is an SAT of bacterial origin.

8. Method according to claim 6, characterized in that the SAT is a plant cytoplasmic SAT, in particular from *Arabidopsis thaliana*.

9. Method according to claim 8, characterized in that the SAT is SAT3 which is represented by SEQ ID NO 1.

10. Method according to claim 6, characterized in that the SAT is a non-cytoplasmic plant SAT from which has been removed its signal(s) for addressing to cellular compartments other than the cytoplasm.

11. Method according to claim 10, characterized in that the SAT is SAT1' which is represented by SEQ ID NO 2.

12. Method according to one of claims 1 to 5, characterized in that the SAT is overexpressed in mitochondria.

13. Method according to claim 12, characterized in that the SAT is overexpressed in the cytoplasm in the form of a signal peptide/SAT fusion protein, the mature functional SAT being released inside mitochondria.

14. Method according to claim 13, characterized in that the mitochondrial addressing signal peptide consists of at least one signal peptide from a natural plant protein which is located in mitochondria, such as for example, the SAT1 signal

peptide which is represented by amino acids 1 to 63 in SEQ ID NO 3.

15. Method according to claim 13, characterized in that the SAT is a mitochondrial SAT of plant origin, in particular from *Arabidopsis thaliana*.

16. Method according to claim 15, characterized in that the SAT is SAT1 which is represented by SEQ ID NO 3.

17. Method according to claim 6, characterized in that the SAT is overexpressed in chloroplasts of plant cells.

18. Method according to claim 17, characterized in that the SAT is overexpressed in chloroplasts by integration, into chloroplast DNA of plant cells, of a chimeric gene comprising a DNA sequence encoding the said SAT, under the control of 5′ and of 3′ regulatory elements which are functional in chloroplasts.

19. Method according to claim 17, characterized in that the SAT is overexpressed in the cytoplasm in the form of a transit peptide/SAT fusion protein, the mature functional SAT being released inside chloroplasts.

20. Method according to claim 19, characterized in that the SAT is homologous with the transit peptide.

21. Method according to claim 20, characterized in that the SAT is a chloroplast SAT of plant origin, in particular from *Arabidopsis thaliana*.

22. Method according to claim 21, characterized in that the SAT is SAT2 or SAT4 which are represented by SEQ ID NO 5 or NO 6, respectively.

23. Method according to claim 19, characterized in that the SAT is heterologous with the transit peptide.

24. Method according to claim 13, characterized in that the SAT is a cytoplasmic SAT of plant origin or an SAT of bacterial origin, as defined in one of claims 3 to 5 or 9 to 11.

25. Method according to either of claims 23 and 24, characterized in that the transit peptide is a transit peptide from another protein which is located in plastids.

26. Method according to claim 25, characterized in that the transit peptide consists of a plant EPSPS transit peptide or a plant RuBisCO ssu transit peptide.

27. Method according to either of claims 25 and 26, characterized in that the transit peptide comprises a transit peptide from a plant protein which is located in plastids, and, between the C-terminal portion of the transit peptide and the N-terminal portion of the SAT, a portion of sequence from the

mature N-terminal region of a protein which is located in plastids.

28. Method according to claim 27, characterized in that the portion of sequence comprises generally less than 40 amino acids from the N-terminal portion of the mature protein, preferably less than 30 amino acids, more preferably between 15 and 25 amino acids.

29. Method according to either of claims 27 and 28, characterized in that the transit peptide comprises, between the C-terminal portion of the N-terminal portion of the mature protein and the N-terminal portion of the SAT, a second transit peptide from a plant protein which is located in plastids.

30. Method according to claim 29, characterized in that the transit peptide is an optimized transit peptide (OTP) made by fusing a first transit peptide with a portion of sequence from the mature N-terminal region of a protein which is located in plastids, which is fused with a second transit peptide.

31. Transit peptide/SAT fusion protein, characterized in that the SAT is heterologous with the transit peptide.

32. Fusion protein according to claim 31, as defined in claims 24 to 30.

33. Nucleic acid sequence encoding a transit peptide/SAT fusion protein according to either of claims 31 and 32.

34. Chimeric gene comprising a coding sequence as well as heterologous 5' and 3' regulatory sequences, which are able to function in a host organism, characterized in that the coding sequence comprises at least one nucleic acid sequence which encodes an SAT.

35. Chimeric gene according to claim 34, characterized in that the host organism is chosen from bacteria, for example *E. coli*, yeasts, in particular of the genera *Saccharomyces*, *Kluyveromyces* or *Pichia*, fungi, in particular *Aspergillus*, baculoviruses, or plant cells and plants.

36. Chimeric gene according to claim 35, characterized in that the host organism is a plant cell or a plant which contains it .

37. Chimeric gene according to claim 36, characterized in that the 5' regulatory element comprises regulatory sequences which are promoters in plant cells and plants, and are chosen from promoters which are expressed in plant leaves, constitutive promoters, or light-dependent promoters of bacterial, viral or plant origin.

38. Chimeric gene according to claim 36, characterized in that the 5' regulatory element comprises regulatory sequences which are promoters in

plant cells and plants, and are chosen from seed-specific promoters.

39. Chimeric gene according to claim 38, characterized in that the promoter is chosen from the promoters for napin, phaseolin, glutenin, zein, helianthinin, albumin and oleosin.

40. Chimeric gene according to one of claims 34 to 39, characterized in that the nucleic acid sequence which encodes an SAT encodes an SAT as defined in claims 2 to 30.

41. Chimeric gene according to one of claims 34 to 39, characterized in that the nucleic acid sequence which encodes an SAT is the nucleic acid sequence according to claim 33.

42. Cloning and/or expression vector for transforming a host organism, characterized in that it contains at least one chimeric gene as defined according to one of claims 34 to 41.

43. Method of transforming host organisms, characterized in that at least one nucleic acid sequence according to claim 33, or a chimeric gene according to one of claims 34 to 41, is integrated into the genome of the said host organism.

44. Method according to claim 43, by means of the vector according to claim 42.

45. Method according to either of claims 43 and 44, characterized in that the host organism is chosen from bacteria, for example *E. coli*, yeasts, in

particular of the genera *Saccharomyces, Kluyveromyces* or *Pichia*, fungi, in particular *Aspergillus*, baculoviruses, or plant cells and plants.

46. Method according to claim 45, characterized in that the host organism is a plant cell or a plant which contains it.

47. Method according to claim 46, characterized in that the plant is regenerated from a transformed plant cell.

48. Method according to claim 47, characterized in that the host organism is a monocotyledonous plant, in particular chosen from cereals, sugar cane, rice and maize, or a dicotyledonous plant, in particular chosen from tobacco, soybean, rape, cotton, beet and clover.

49. Transformed host organism, characterized in that it comprises at least one nucleic acid sequence according to claim 33, or a chimeric gene according to one of claims 34 to 41.

50. Host organism according to claim 49, characterized in that it is obtained by the method according to one of claims 43 to 48.

51. Plant cell, characterized in that it comprises at least one nucleic acid sequence according to claim 33, or a chimeric gene according to one of claims 34 to 41.

52. Genetically modified plant, characterized in that it comprises at least one plant cell according to claim 51.

53. Plant according to claim 52, characterized in that the plant is regenerated from a plant cell according to claim 51.

54. Genetically modified plant, characterized in that it is derived from the culture and/or crossing of regenerated plants, according to claim 53.

55. Genetically modified plant according to one of claims 52 to 54, characterized in that it is a monocotyledonous plant, in particular chosen from cereals, sugar cane, rice and maize, or a dicotyledonous plant, in particular chosen from tobacco, soybean, rape, cotton, beet and clover.

56. Genetically modified plant according to one of claims 52 to 55, characterized in that it comprises other genes of interest.

57. Genetically modified plant according to claim 56, characterized in that it comprises at least one other gene which modifies the content and quality of the proteins of the said plant, in particular in the leaves and/or seeds.

58. Genetically modified plant according to either of claims 56 and 57, characterized in that the gene encodes a protein enriched in sulphur-containing amino acids.

59. Seeds of genetically modified plants according to one of claims 52 to 58.

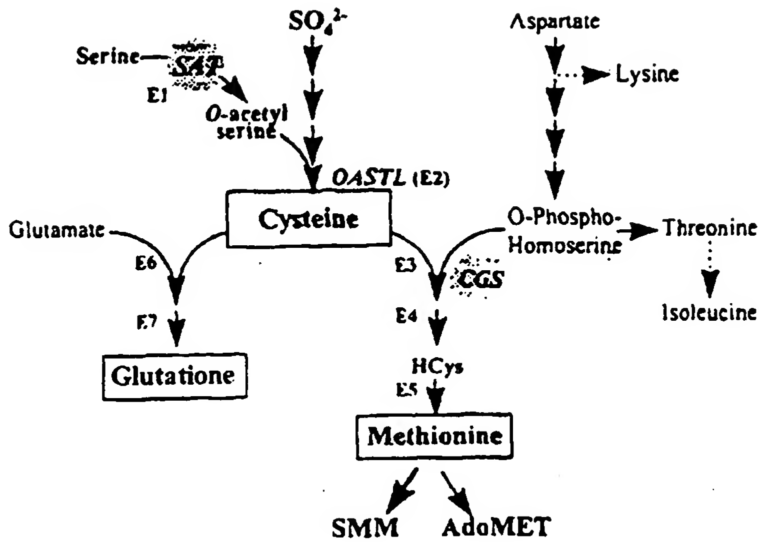


Figure 1: Sequence illustrating the synthetic pathway for cysteine and sulphur derivatives (glutathione and methionine)

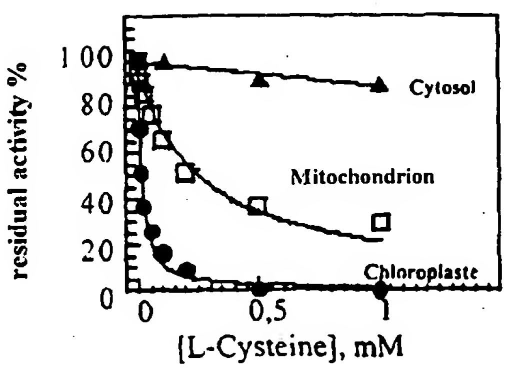


Figure 2: Effect of cysteine on serine acetyltransferase activities in pea (*Pisum sativum*)

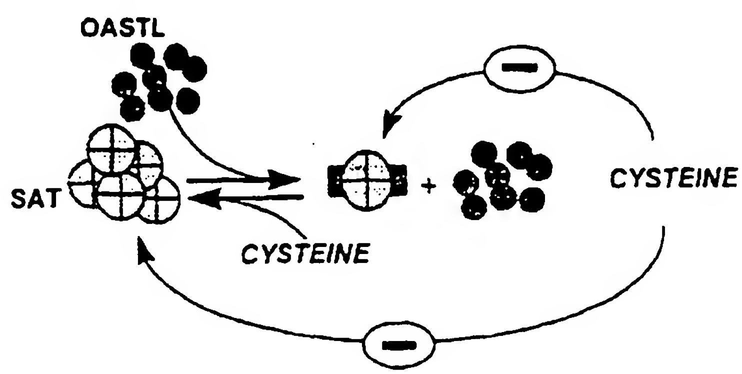


Figure 3: Model of inhibition of chloroplast serine acetyltransferase

M	A	T	C	I	D	T	C	R	T	G	N	T	Q	D	D	16	
ATG	**GCA**	**ACA**	**TGC**	**ATA**	**GAC**	**ACA**	**TGC**	CGA	ACC	GGT	AAT	ACC	CAA	GAC	GAT		48
D	S	R	F	C	C	I	K	N	F	F	R	P	G	F	S	32	
GAT	TCC	CGG	TTC	TGT	TGC	ATC	AAG	AAT	TTC	TTT	CGA	CCC	GGT	TTC	TCT		96
V	N	R	K	I	H	H	T	Q	I	E	D	D	D	D	V	48	
GTA	AAC	CGG	AAG	ATT	CAC	CAC	ACC	CAA	ATC	GAA	GAT	GAC	GAT	GAT	GTC		144
W	I	K	M	L	E	E	A	K	S	D	V	K	Q	E	P	64	
TGG	ATC	AAG	ATG	CTT	GAA	GAA	GCC	AAA	TCC	GAT	GTT	AAA	CAA	GAA	CCC		192
I	L	S	N	Y	Y	Y	A	S	I	T	S	H	R	S	L	80	
ATT	TTA	TCA	AAC	TAC	TAC	TAC	GCT	TCG	ATC	ACA	TCT	CAT	CGA	TCT	TTA		240
E	S	A	L	A	H	I	L	S	V	K	L	S	N	L	N	96	
GAG	TCT	GCT	TTA	GCT	CAC	ATC	CTC	TCC	GTA	AAG	CTC	AGC	AAT	TTA	AAC		288
L	P	S	N	T	L	F	E	L	F	I	S	V	L	E	E	112	
CTA	CCA	AGC	AAC	ACA	CTC	TTC	GAA	CTG	TTC	ATA	AGC	GTT	TTA	GAA	GAA		336
S	P	E	I	I	E	S	T	K	Q	D	L	I	A	V	K	128	
AGC	CCT	GAG	ATC	ATC	GAA	TCC	ACG	AAG	CAA	GAT	CTT	ATA	GCA	GTC	AAA		384
E	R	D	P	A	C	I	S	Y	V	H	C	F	L	G	F	144	
GAA	AGA	GAC	CCA	GCT	TGT	ATA	AGC	TAC	GTT	CAT	TGC	TTC	TTG	GGC	TTC		432
K	G	F	L	A	C	Q	A	H	R	I	A	H	T	L	W	160	
AAA	GGC	TTC	CTC	GCT	TGT	CAA	GCT	CAT	CGA	ATA	GCT	CAT	ACC	CTC	TGG		480
K	Q	N	R	K	I	V	A	L	L	I	Q	N	R	V	S	176	
AAA	CAG	AAC	AGA	AAA	ATC	GTA	GCT	TTA	TTG	ATC	CAA	AAC	AGA	GTA	TCA		528
E	S	F	A	V	D	I	H	P	G	A	K	I	G	K	G	192	
GAA	TCT	TTC	GCC	GTC	GAT	ATT	CAT	CCC	GGA	GCG	AAG	ATC	GGA	AAA	GGG		576
I	L	L	D	H	A	T	G	V	V	I	G	E	T	A	V	208	
ATT	CTT	TTA	GAC	CAT	GCG	ACG	GGC	GTG	GTG	ATC	GGA	GAG	ACG	GCG	GTG		624
V	G	D	N	V	S	I	L	H	G	V	T	L	G	G	T	224	
GTT	GGA	GAC	AAT	GTT	TCG	ATT	CTA	CAC	GGA	GTG	ACC	TTG	GGA	GGA	ACA		672
G	K	Q	S	G	D	R	H	P	K	I	G	D	G	V	L	240	
GGG	AAA	CAG	AGT	GGT	GAT	CGG	CAT	CCG	AAG	ATT	GGT	GAT	GGT	GTG	TTG		720
I	G	A	G	S	C	I	L	G	N	I	T	I	G	E	G	256	
ATT	GGA	GCT	GGG	AGT	TGT	ATA	TTG	GGG	AAT	ATA	ACA	ATC	GGT	GAG	GGA		768
A	K	I	G	S	G	S	V	V	V	K	D	V	P	A	R	272	
GCT	AAG	ATT	GGA	TCA	GGG	TCG	GTG	GTG	GTT	AAG	GAT	GTG	CCG	GCG	CGT		816
T	T	A	V	G	N	P	A	R	L	I	G	G	K	E	N	288	
ACG	ACG	GCG	GTT	GGA	AAT	CCG	GCG	AGG	TTG	ATT	GGT	GGG	AAA	GAG	AAT		864
P	R	K	H	D	K	I	P	C	L	T	M	D	Q	T	S	304	
CCG	AGA	AAA	CAT	GAT	AAG	ATT	CCT	TGT	CTG	ACT	ATG	GAC	CAG	ACA	TCG		912
Y	L	T	E	W	S	D	Y	V	I							314	
TAT	TTA	ACC	GAG	TGG	TCT	GAT	TAT	GTG	ATT	TAA							945

Figure 4: Nucleotide and peptide sequence of the SAT3 isoform gene (L34076) of *A. thaliana*

		M	P	P	A	G	E	L	R	H	Q	S	P	S	K	14
		ATG	CCA	CCG	GCC	GGA	GAA	CTC	CGA	CAT	CAA	TCT	CCA	TCA	AAG	42
E	K	L	S	S	V	T	Q	S	D	E	A	E	A	A	S	30
GAG	AAA	CTA	TCT	TCC	GTT	ACC	CAA	TCC	GAT	GAA	GCA	GAA	GCA	GCG	TCA	90
A	A	I	S	A	A	A	A	D	A	E	A	A	G	L	W	46
GCA	GCG	ATA	TCT	GCG	GCA	GCT	GCA	GAT	GCG	GAA	GCT	GCC	GGA	TTA	TGG	138
T	Q	I	K	A	E	A	R	R	D	A	E	A	E	P	A	62
ACA	CAG	ATC	AAG	GCG	GAA	GCT	CGC	CGT	GAT	GCT	GAG	GCG	GAG	CCA	GCT	186
L	A	S	Y	L	Y	S	T	I	L	S	H	S	S	L	E	78
TTA	GCT	AGC	TAT	CTA	TAT	TCG	ACG	ATT	CTT	TCT	CAT	TCG	TCT	CTT	GAA	234
R	S	I	S	F	H	L	G	N	K	L	C	S	S	T	L	94
CGA	TCT	ATC	TCG	TTT	CAT	CTA	GGA	AAC	AAG	CTT	TGT	TCC	TCA	ACG	CTT	282
L	S	T	L	L	Y	D	L	F	L	N	T	F	S	S	D	110
TTA	TCC	ACA	CTT	TTA	TAC	GAT	CTG	TTC	TTA	AAC	ACT	TTT	TCC	TCC	GAT	330
P	S	L	R	N	A	T	V	A	D	L	R	A	A	R	V	126
CCT	TCT	CTT	CGT	AAC	GCC	ACC	GTC	GCA	GAT	CTA	CGC	GCT	GCT	CGT	CTT	378
R	D	P	A	C	I	S	F	S	H	C	L	L	N	Y	K	142
CGT	GAT	CCT	GCT	TGT	ATC	TCG	TTC	TCT	CAT	TGT	CTC	CTC	AAT	TAC	AAA	426
G	F	L	A	I	Q	A	H	R	V	S	H	K	L	W	T	158
GGC	TTC	TTA	GCT	ATT	CAG	GCG	CAT	CGT	GTA	TCA	CAC	AAC	CTA	TGG	ACA	474
Q	S	R	K	P	L	A	L	[illegible]	L	[illegible]	[illegible]	R	[illegible]	[illegible]	[illegible]	[illegible]
CAA	TCA	CGG	AAG	CCA	TTA	GCA	TTA	GTT	CTA	CAC	TTA	AGA	ATC	TCC	GAT	522
V	F	A	V	D	I	H	P	A	A	K	I	G	K	G	I	190
GTA	TTC	GCT	GTT	GAT	ATC	CAT	CCA	GCA	GCG	AAG	ATC	GGA	AAA	GGG	ATA	570
L	L	D	H	A	T	G	V	V	V	G	E	T	A	V	I	206
CTT	CTA	GAC	CAC	GCA	ACC	GGA	GTT	GTA	GTC	GGA	GAA	ACA	GCG	GTG	ATT	618
G	N	N	V	S	I	L	H	H	V	T	L	G	G	T	G	222
GGG	AAC	AAT	GTT	TCA	ATC	CTT	CAC	CAT	GTG	ACA	CTA	GGT	GGA	ACA	GGT	666
K	A	C	G	D	R	H	P	K	I	G	D	G	C	L	I	238
AAA	GCT	TGT	GGA	GAT	AGA	CAT	CCG	AAG	ATC	GGT	GAC	GGT	TGT	TTG	ATT	714
C	A	G	A	T	I	L	G	N	V	K	I	G	A	G	A	254
GGA	GCT	GGA	GCG	ACT	ATT	CTT	GGA	AAT	GTG	AAG	ATT	GGT	GCA	GGT	GCT	762
K	V	G	A	G	S	V	V	L	I	D	V	P	C	R	G	270
AAA	GTA	GGA	GCT	GGT	TCT	GTT	GTG	CTG	ATT	GAC	GTG	CCT	TGT	CGA	GGT	810
T	A	V	G	N	P	A	R	L	V	G	G	K	E	K	P	286
ACT	CCG	GTT	GGG	AAT	CCG	GCG	AGA	CTT	GTC	GGA	GGG	AAA	GAG	AAG	CCA	858
T	I	H	D	E	E	C	P	G	E	S	M	D	H	T	S	302
ACG	ATT	CAT	GAT	GAG	GAA	TGT	CCT	GGA	GAA	TCG	ATG	GAT	CAT	ACT	TCA	906
F	I	S	E	W	S	D	Y	I	I...							312
TTC	ATC	TCG	GAA	TGG	TCA	GAT	TAC	ATC	ATA	TAA						939

Figure 5: Nucleotide and peptide sequence of the SAT3' isoform gene (U30298) of *A. thaliana*

```
M   A   A   C   I   D   T   C   R   T   G   K   P   Q   I           15
ATG GCT GCG TGC ATC GAC ACC TGC CGC ACT GGT AAA CCC CAG ATT         45
S   P   R   D   S   S   K   H   H   D   D   E   S   G   F           30
TCT CCT CGC GAT TCT TCT AAA CAC CAC GAC GAT GAA TCT GGC TTT         90
R   Y   M   N   Y   F   R   Y   P   D   R   S   S   F   N           45
CGT TAC ATG AAC TAC TTC CGT TAT CCT GAT CGA TCT TCC TTC AAT        135
G   T   Q   T   K   T   L   H   T   R   P   L   L   E   D           60
GGA ACC CAG ACC AAA ACC CTC CAT ACT CGT CCT TTG CTT GAA GAT        180
L   D   R   D   A   E   V   D   D   V   W   A   K   I   R           75
CTC GAT CGC GAC GCT GAA GTC GAT GAT GTT TGG GCC AAA ATC CGA        225
E   E   A   K   S   D   I   A   K   E   P   I   V   S   A           90
GAA GAG GCT AAA TCT GAT ATC GCC AAA GAA CCT ATT GTT TCC GCT        270
Y   Y   H   A   S   I   V   S   Q   R   S   L   E   A   A          105
TAT TAT CAC GCT TCG ATT GTT TCT CAG CGT TCG TTG GAA GCT GCG        315
L   A   N   T   L   S   V   K   L   S   N   L   N   L   P          120
TTG GCG AAT ACT TTA TCT GTT AAA CTC AGC AAT TTG AAT CTT CCA        360
S   N   T   L   F   D   L   F   S   G   V   L   Q   G   N          135
AGC AAC ACG CTT TTC GAT TTG TTC TCT GGT GTT CTT CAA GGA AAC        405
P   D   I   V   E   S   V   K   L   D   L   L   A   V   K          150
CCA GAT ATT GTT GAA TCT GTC AAG CTA GAT CTT TTA GCT GTT AAG        450
E   R   D   P   A   C   I   S   Y   V   H   C   F   L   H          165
GAG AGA GAT CCT GCT TGT ATA AGC TAC GTT CAT TGT TTC CTT CAT        495
F   K   G   F   L   A   C   Q   A   H   R   I   A   H   E          180
TTT AAA GGC TTC CTC GCT TGT CAA GCG CAT CGT ATT GCT CAT GAG        540
L   W   T   Q   D   R   K   I   L   A   L   L   I   Q   N          195
CTT TGG ACT CAG GAC AGA AAA ATC CTA GCT TTG TTG ATC CAG AAC        585
R   V   S   E   A   F   A   V   D   F   H   P   G   A   K          210
AGA GTC TCT GAA GCC TTC GCT GTT GAT TTC CAC CCT GGA GCT AAA        630
I   G   T   G   I   L   L   D   H   A   T   A   I   V   I          225
ATC GGT ACC GGG ATT TTG CTA GAC CAT GCT ACG GCT ATT GTG ATC        675
G   E   T   A   V   V   G   N   N   V   S   I   L   H   N          240
GGT GAG ACG GCG GTT GTG GGG AAC AAT GTT TCG ATT CTC CAT AAC        720
V   T   L   G   G   T   G   K   Q   C   G   D   R   H   P          255
GTT ACG CTT GGA GGA ACG GGG AAA CAG TGT GGA GAT AGG CAC CCG        765
K   I   G   D   G   V   L   I   G   A   G   T   C   I   L          270
AAG ATT GGC GAT GGG GTT TTG ATT GGA GCT GGG ACT TGT ATT TTG        810
G   N   I   T   I   G   E   G   A   K   I   G   A   G   S          285
GGG AAT ATC ACG ATT GGT GAA GGA GCT AAG ATT GGT GCG GGG TCG        855
V   V   L   K   D   V   P   P   R   T   T   A   V   G   N          300
GTG GTG TTG AAA GAC GTG CCG CCG CGT ACG ACG GCT GTT GGA AAT        900
P   A   R   L   L   G   G   K   D   N   P   K   T   H   D          315
CCG GCG AGG TTG CTT GGT GGT AAA GAT AAT CCG AAA ACG CAT GAC        945
K   I   P   G   L   T   M   D   Q   T   S   H   I   S   E          330
AAG ATT CCT GGT TTG ACT ATG GAC CAG ACG TCG CAT ATA TCC GAG        990
W   S   D   Y   V   I                                              336
TGG TCG GAT TAT GTA ATT TGA                                  1011
```

Figure 6: Nucleotide and peptide sequence of a SAT1' isoform gene (L78443) of *A. thaliana*

```
                              M   L   P   V   T   S   R   R   H   F              10
                              ATG TTG CCG GTC ACA AGT CGC CGC CAC TTC      30
T   M   S   L   Y   M   L   R   S   S   S   P   H   I   N                        25
ACA ATG TCC CTA TAT ATG CTC CGT TCA TCT TCT CCA CAC ATC AAT                   75
H   H   S   F   L   L   P   S   F   V   S   S   K   F   K                        40
CAT CAC TCT TTC CTT CTT CCT TCT TTT GTT TCC TCC AAA TTC AAA                  120
H   H   T   L   S   P   P   P   S   P   P   P   P   P   P                        55
CAC CAT ACT TTA TCT CCT CCT CCT TCT CCT CCT CCT CCT CCT CCT                  165
M   A   A   C   I   D   T   C   R   T   G   K   P   Q   I                        70
ATG GCT GCG TGC ATC GAC ACC TGC CGC ACT GGT AAA CCC CAG ATT                  210
S   P   R   D   S   S   K   H   H   D   D   E   S   G   F                        85
TCT CCT CGC GAT TCT TCT AAA CAC CAC GAC GAT GAA TCT GGC TTT                  255
R   Y   M   N   Y   F   R   Y   P   D   R   S   S   F   N                       100
CGT TAC ATG AAC TAC TTC CGT TAT CCT GAT CGA TCT TCC TTC AAT                  300
G   T   Q   T   K   T   L   H   T   R   P   L   L   E   D                       115
GGA ACC CAG ACC AAA ACC CTC CAT ACT CGT CCT TTG CTT GAA GAT                  345
L   D   R   D   A   E   V   D   D   V   W   A   K   I   R                       130
CTC GAT CGC GAC GCT GAA GTC CAT GAT GTT TGG GCC AAA ATC CGA                  390
E   E   A   K   S   D   I   A   K   E   P   I   V   S   A                       145
GAA GAG GCT AAA TCT GAT ATC GCC AAA GAA CCT ATT GTT TCC GCT                  435
Y   Y   H   A   S   I   V   S   Q   R   S   L   E   A   A                       160
TAT TAT CAC GCT TCG ATT GTT TCT CAG CGT TCG TTG GAA GCT GCG                  480
L   A   N   T   L   S   V   K   L   S   N   L   N   L   P                       175
TTG GCG AAT ACT TTA TCT GTT AAA CTC AGC AAT TTG AAT CTT CCA                  525
S   N   T   L   F   D   L   F   S   G   V   L   Q   G   N                       190
AGC AAC ACG CTT TTC GAT TTG TTC TCT GGT GTT CTT CAA GGA AAC                  570
P   D   I   V   E   S   V   K   L   D   L   L   A   V   K                       205
CCA CAT ATT GTT GAA TCT GTC AAG CTA GAT CTT TTA GCT GTT AAC                  615
E   R   D   P   A   C   I   S   Y   V   H   C   F   L   H                       220
GAG AGA GAT CCT GCT TGT ATA AGC TAC GTT CAT TGT TTC CTT CAC                  660
F   K   G   F   L   A   C   Q   A   H   R   I   A   H   E                       235
TTT AAA GGC TTC CTC GCT TGT CAA GCG CAT CGT ATT GCT CAT GAG                  705
L   W   T   Q   D   R   K   I   L   A   L   L   I   Q   N                       250
CTT TGG ACT CAG GAC AGA AAA ATC CTA GCT TTG TTG ATC CAG AAC                  750
R   V   S   E   A   F   A   V   D   F   H   P   G   A   K                       265
AGA GTC TCT GAA GCC TTC GCT GTT GAT TTC CAC CCT GGA GCT AAA                  795
I   G   T   G   I   L   L   D   H   A   T   A   I   V   I                       280
ATC GGT ACC GGG ATT TTG CTA GAC CAT GCT ACG GCT ATT GTG ATC                  840
G   E   T   A   V   V   G   N   N   V   S   I   L   H   N                       295
GGT GAG ACG GCG GTT GTG GGG AAC AAT GTT TCG ATT CTC CAT AAC                  885
V   T   L   G   G   T   G   K   Q   C   G   D   R   H   P                       310
GTT ACG CTT GGA GGA ACG GGG AAA CAG TGT GGA GAT AGG CAC CCG                  930
K   I   G   D   G   V   L   I   G   A   G   T   C   I   L                       325
AAG ATT GGC GAT GGG GTT TTG ATT GGA GCT GGG ACT TGT ATT TTG                  975
G   N   I   T   I   G   E   G   A   K   I   G   A   G   S                       340
GGG AAT ATC ACG ATT GGT GAA GGA GCT AAG ATT GGT GCG GGG TCG                 1020
V   V   L   K   D   V   P   P   R   T   T   A   V   G   N                       355
GTG GTG TTG AAA GAC GTG CCG CCG CGT ACG ACG GCT GTT GGA AAT                 1065
P   A   R   L   L   G   G   K   D   N   P   K   T   H   D                       370
CCG GCG AGG TTG CTT GGT GGT AAA GAT AAT CCG AAA ACG CAT GAC                 1110
K   I   P   G   L   T   M   D   Q   T   S   H   I   S   E                       385
AAG ATT CCT GGT TTG ACT ATG GAC CAG ACG TCG CAT ATA TCC GAG                 1155
W   S   D   Y   V   I                                                        391
TGG TCG GAT TAT GTA ATT TGA                                          1176
```

Figure 7: Nucleotide and peptide sequence of an SAT1 isoform gene (U22964) of *A. thaliana*

```
 M   V   D   L   S   S   F   S   L   L   F   A   F   S   V   S       16
ATG GTG GAT CTA TCT TCC TTT AGC CTC CTT TTT GCT TTC TCC GTC TCT          48
 L   S   F   V   Q   S   K   R   V   C   D   S   S   L   S   S       32
CTC TCT TTT GTC CAA TCA AAA AGA GTT TGT GAT TCT TCT TTA TCG TCT          96
 P   W   R   D   M   N   G   D   E   L   P   F   E   S   G   F       48
CCT TGG AGA GAT ATG AAT GGC GAT GAG CTT CCT TTC GAG AGT GGT TTC         144
 E   V   Y   A   K   G   T   H   K   S   E   F   D   S   N   L       64
GAG GTT TAC GCT AAG GGA ACT CAT AAG TCA GAG TTT GAC TCG AAT TTG         192
 L   D   P   R   S   D   P   I   W   D   A   I   R   E   E   A       80
CTT GAT CCT CGT TCT GAT CCT ATT TGG GAT GCT ATA AGA GAA GAA GCT         240
 K   L   E   A   E   K   E   P   I   L   S   S   F   L   Y   A       96
AAA CTT GAG GCA GAG AAA GAG CCT ATT TTG AGT AGC TTC TTG TAT GCT         288
 G   I   L   A   H   D   C   L   E   Q   A   L   G   F   V   L      112
GGT ATC TTA GCA CAT GAT TGT TTA GAG CAA GCT TTA GGG TTT GTT CTA         336
 A   N   R   L   Q   N   P   T   L   L   A   T   Q   L   L   D      128
GCC AAC CGT CTC CAA AAC CCA ACC TTG TTG GCA ACA CAA CTC TTG GAT         384
 I   F   Y   G   V   M   M   H   D   K   G   I   Q   S   S   I      144
ATA TTT TAT GGT GTT ATG ATG CAT GAC AAA GGT ATT CAG AGT TCG ATT         432
 R   H   D   L   Q   A   F   K   D   R   D   F   A   C   L   S      160
CGG CAT GAT CTC CAG GCA TTT AAA GAT CGT GAT TTT GCT TGT CTG TCG         480
 Y   S   S   A   I   L   H   L   K   G   Y   H   A   L   Q   A      176
TAT AGT TCT GCT ATT TTA CAT CTG AAA GGT TAT CAT GCG TTA CAA GCA         528
 Y   R   V   A   H   K   L   W   N   E   G   R   K   L   L   A      192
TAT AGG GTT GCG CAT AAA CTG TGG AAT GAA GGG AGG AAA CTA TTA GCT         576
 L   A   L   Q   S   R   I   S   E   V   F   G   I   D   I   H      208
CTT GCA TTG CAA AGC CGA ATA AGC GAG GTT TTT GGC ATT GAC ATA CAT         624
 P   A   A   K   I   G   E   G   I   L   L   D   H   G   T   G      224
CCA GCG GCA AGA ATT GGG GAG GGA ATA TTG TTG GAT CAT GGA ACT GGA         672
 V   V   I   G   E   T   A   V   I   G   N   G   V   S   I   L      240
GTG GTC ATT GGT GAG ACC GCT GTG ATA GGC AAC GGT GTC TCG ATC TTA         720
 H   G   V   T   L   G   G   T   G   K   E   T   G   D   R   H      256
CAT GGT GTG ACT TTA GGA GGA ACC GGA AAG GAA ACT GGC GAT CGC CAC         768
 P   K   I   G   E   G   A   L   L   G   A   C   V   T   I   L      272
CCA AAG ATA GGT GAA GGT GCA TTG CTT GGA GCT TGT GTG ACT ATA CTT         816
 G   N   I   S   I   G   A   G   A   M   V   A   A   G   S   L      288
GGT AAC ATA AGC ATA GGT GCT GGA GCA ATG GTA GCT GCA GGT TCA CTT         864
 V   L   K   D   V   P   S   H   S   V   V   A   G   N   P   A      304
GTG TTA AAA GAC GTT CCT TCG CAT AGT GTG GTG GCT GGA AAT CCT GCA         912
 K   L   I   R   V   M   E   E   Q   D   P   S   L   A   M   K      320
AAA CTG ATC AGG GTC ATG GAA GAG CAA GAC CCG TCT CTA GCA ATG AAA         960
 H   D   A   T   K   E   F   F   R   H   V   A   D   G   Y   K      336
CAC GAT GCT ACT AAA GAG TTC TTT CGA CAT GTA GCT GAT GGT TAC AAA        1008
 G   A   Q   S   N   G   P   S   L   S   A   G   D   T   E   K      352
GGG GCA CAA TCT AAC GGA CCA TCA CTT TCA GCA GGA GAT ACA GAG AAA        1056
 G   H   T   N   S   T   S                                          359
GGA CAC ACT AAC AGC ACA TCA TGA                                        1104
```

Figure 8: Nucleotide and peptide sequence of the mRNA of the putative chloroplastic serine acetyltransferase SAT2 of *Arabidopsis thaliana* (L78444)

```
M   A   C   I   N   G   E   N   R   D   F   S   S   S   S        15
ATG GCT TGT ATA AAC GGC GAG AAT CGT GAT TTT TCT TCC TCG TCA           45
S   L   S   S   L   P   M   I   V   S   R   N   F   S   A        30
TCT TTG TCT TCT CTT CCA ATG ATT GTC TCC CGG AAC TTT TCT GCC           90
R   D   D   G   E   T   G   D   E   F   P   F   E   R   I        45
AGA GAC GAT GGA GAG ACC GGT GAC GAG TTT CCT TTC GAG AGG ATT          135
F   P   V   Y   A   R   G   T   L   N   P   V   A   D   P        60
TTC CCG GTT TAC GCT AGA GGA ACC CTT AAT CCC GTG GCC GAC CCG          180
V   L   L   D   F   T   N   S   S   Y   D   P   I   W   D        75
GTT TTG CTG GAT TTT ACC AAT TCT AGT TAT GAC CCA ATT TGG GAT          225
S   I   R   E   E   A   K   L   E   A   E   E   E   P   V        90
TCT ATA AGA GAA GAA GCT AAG CTT GAG GCA GAA GAG GAG CCG GTT          270
L   S   S   F   L   Y   A   S   I   L   S   H   D   C   L       105
TTG AGT AGC TTC TTG TAT GCT AGT ATC TTG TCG CAT GAC TGT TTA          315
E   Q   A   L   S   F   V   L   A   N   R   L   Q   N   P       120
GAG CAA GCA TTG AGT TTT GTT CTA GCT AAC CGT CTC CAA AAC CCT          360
T   L   L   A   T   Q   L   M   D   I   F   C   N   V   M       135
ACC TTG TTG GCA ACT CAG CTC ATG GAT ATA TTT TGC AAC GTT ATG          405
V   H   D   R   G   I   Q   S   S   I   R   L   D   V   Q       150
GTA CAT GAC AGA GGT ATT CAA AGC TCG ATT CGT CTT CAT GTT CAG          450
A   F   K   D   R   D   P   A   C   L   S   Y   S   S   A       165
GCA TTC AAA GAC AGA GAT CCT GCT TGT CTA TCG TAT AGT TCG GCT          495
I   L   H   L   K   G   Y   L   A   L   Q   A   Y   R   V       180
ATT TTA CAT CTG AAG GGC TAT CTT GCA CTG CAG GCG TAT AGA GTA          540
A   H   K   L   W   K   Q   G   R   K   L   L   A   L   A       195
GCA CAT AAG TTG TGG AAG CAA GGA AGA AAA CTA TTA GCA TTG GCA          585
L   Q   S   R   V   S   E   V   R   T   A   V   I   G   D       210
CTG CAA AGC CGA GTA AGC GAG GTA AGA ACT GCT GTG ATA GGC GAC          630
R   V   S   I   L   H   G   V   T   L   G   G   T   G   K       225
CGT GTC TCA ATT TTG CAT GGT GTG ACA TTA GGA GGA ACT GGG AAA          675
E   T   G   D   R   H   P   N   I   G   D   G   A   L   L       240
GAA ACC GGT GAC CGC CAT CCA AAT ATA GGC GAC GGT GCT CTT CTT          720
G   A   C   V   T   I   L   G   N   I   K   I   G   A   G       255
GGA GCA TGT GTG ACT ATA CTT GGT AAC ATT AAG ATA GGC GCT GGA          765
A   M   V   A   A   G   S   L   V   L   K   D   V   P   S       270
GCA ATG GTA GCT GCT GGT TCG CTT GTG TTA AAG GAT GTT CCT TCG          810
H   S   M   V   A   G   N   P   A   K   L   I   G   F   V       285
CAT AGC ATG GTG GCT GGA AAT CCA GCA AAA CTC ATC GGG TTT GTT          855
D   E   Q   D   P   S   N   T   M   E   H   G   E   S           299
GAT GAG CAA GAT CCA TCT ATG ACA ATG GAG CAT GGT GAG TCT TGA          900
```

Figure 9: Nucleotide and amino acid sequence of the mRNA of putative chloroplastic SAT4 of *Arabidopsis thaliana*

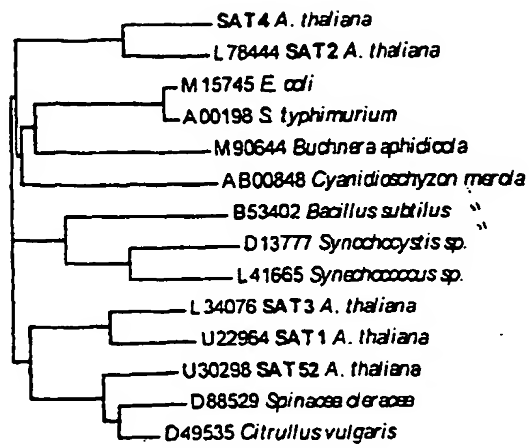


Figure 10: Dendrogram of serine acetyltransferase from several organisms

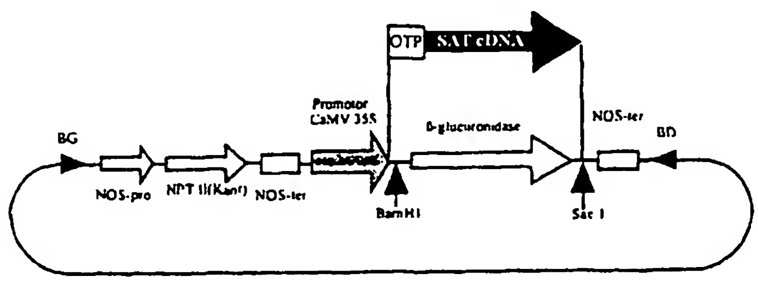


Figure 11: Method of cloning the OTP/Serine acetyltransferase SAT3 or SAT (cysteine-insensitive, for example truncated SAT1) into vector pBI121)

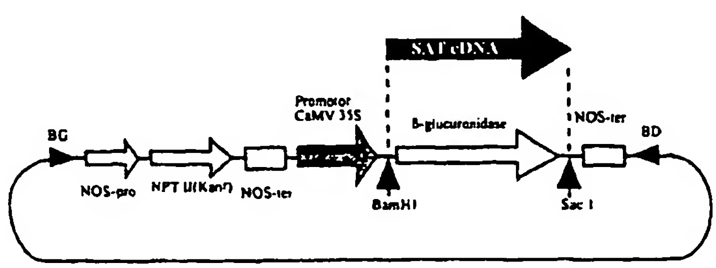


Figure 12: Method for cloning Serine acetyltransferase SAT1'; SAT1; SAT2; SAT3; SAT3'; SAT4, or any SAT into vector pBI121

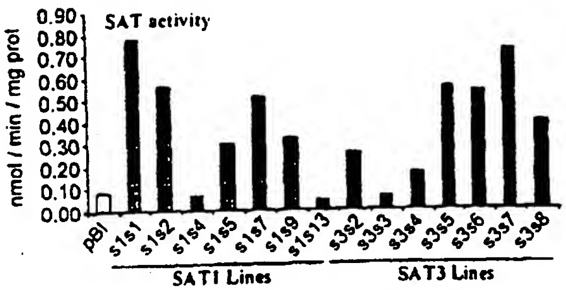


Figure 13

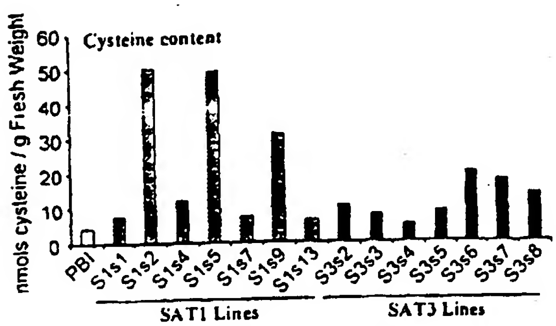


Figure 14

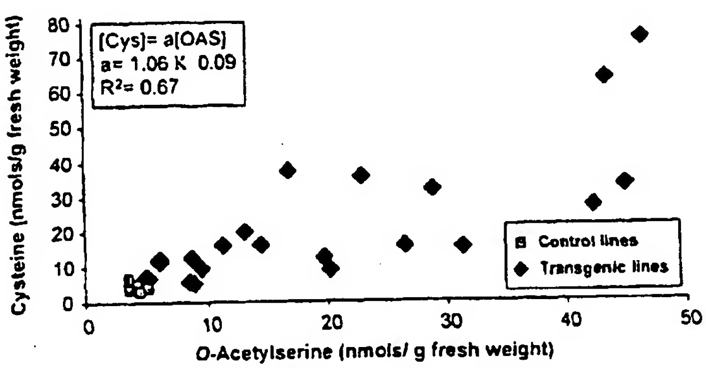


Figure 15

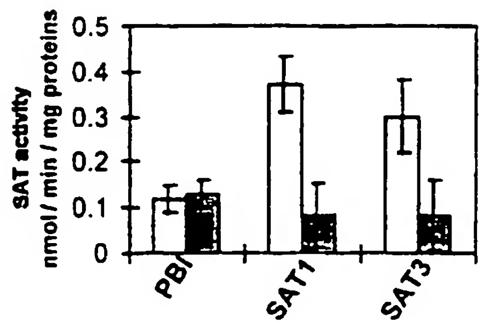


Figure 16

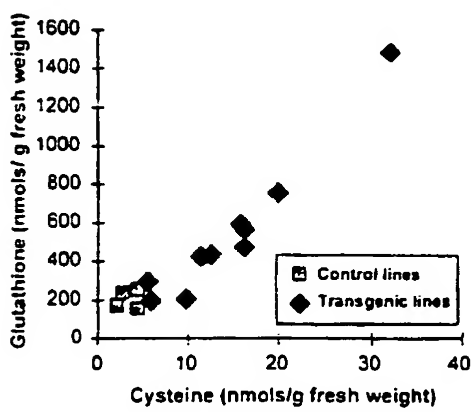


Figure 17

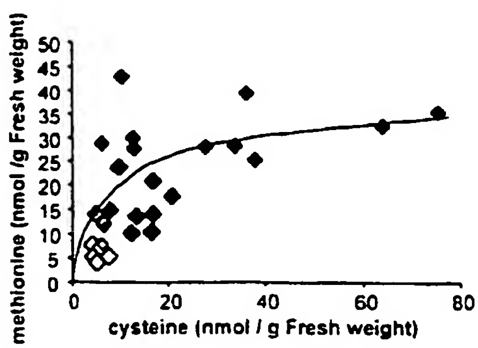


Figure 18

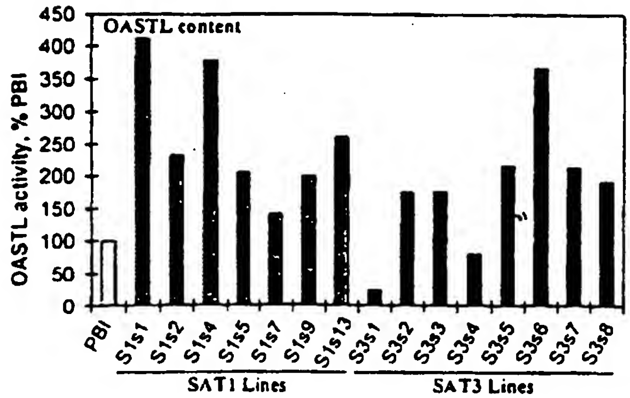


Figure 19

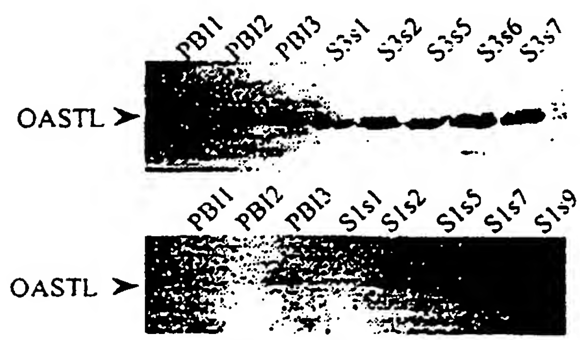


Figure 20

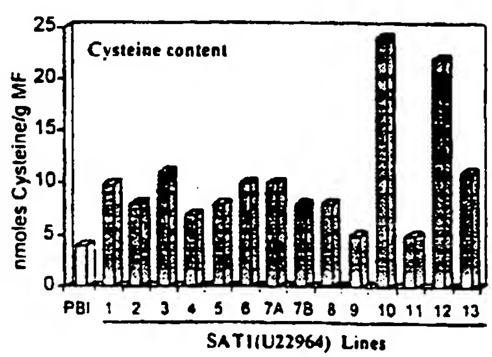


Figure 21

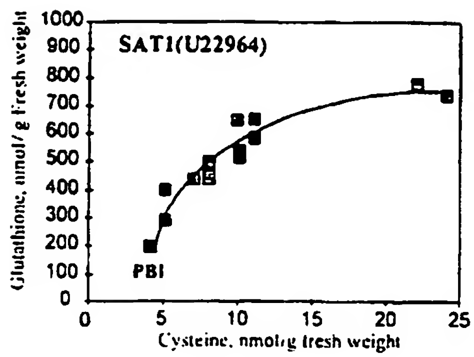
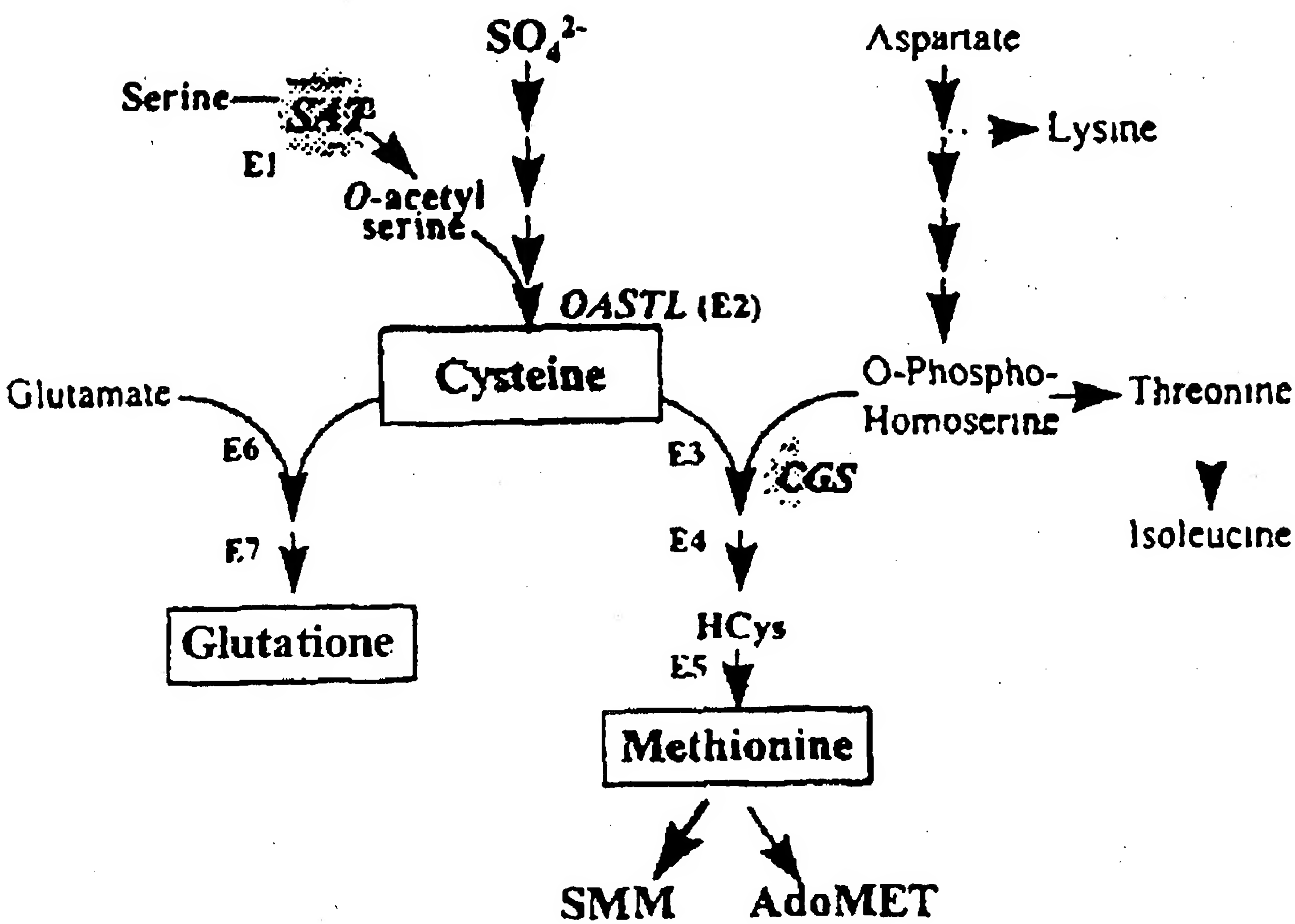

SAT1(U22964)
PBI
Glutathione, nmol/g fresh weight
Cysteine, nmol/g fresh weight
0
100
200
300
400
500
600
700
800
900
1000
0
5
10
15
20
25

Figure 22



Sequence illustrating the synthetic pathway for cysteine and sulphur derivatives (glutathione and methionine)

This Page Blank (uspto)